

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:38:16 ; Search time 23.38 Seconds
(without alignments)
697.011 Million cell updates/sec

Title: US-09-847-519a-2

Perfect score: 1161

Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VQRRRSQDGEEDGREL 220

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_1101.*

- 1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1159	99.8	220	AAE04837	Human SGP003 phosph
2	1026.5	88.4	221	AAE18667	A human regulator
3	465.5	40.1	198	AAE73219	Human phosphatase
4	465.5	40.1	248	AAE68779	Amino acid sequenc
5	465.5	40.1	318	AAE42334	Human ORFX ORF2098
6	463.5	39.9	221	AAE92177	Human cardiovascular
7	461	39.7	198	AAE73214	Murine phosphatase
8	460.5	39.7	327	AAE39977	Human polypeptide
9	456.5	39.3	291	AAE28519	Human DSP-7, Homo
10	445	38.3	211	AAE85620	Human dual specif
11	445	38.3	211	AAE39356	Human polypeptide

12	445	38.3	211	22	AAE04839	Human SGP060 phosph
13	445	38.3	328	22	AAE42355	Human polypeptide
14	443.5	38.2	549	22	AAE04838	Human SGP014 phosph
15	426	36.7	207	22	AAE28793	Human hydrolase-11
16	377	32.5	1133	22	AAE04836	Human SGP018 phosph
17	363	31.3	263	21	AAE92175	Human cardiovascular
18	346.5	29.8	180	22	AAE06787	Human dual-specifi
19	346.5	29.8	180	22	AAE66441	Human MAP-kinase p
20	346.5	29.8	185	15	AAE56968	Human phosphatase
21	346.5	29.8	185	18	AAE35330	Human GDC25B vacci
22	346.5	29.8	185	21	AAE42873	Human ORFX ORF2637
23	302	26.0	150	21	AAE40919	Human ORFX ORF683
24	288	24.8	276	22	AAE39978	Human polypeptide
25	284	24.5	240	21	AAE28520	Human DSP-7 splice
26	256	22.0	167	21	AAE85621	Human dual specif
27	247.5	21.3	118	15	AAE56969	Human phosphatase
28	224	19.3	173	22	AAE41763	Human polypeptide
29	224	19.3	173	22	AAE41764	Human polypeptide
30	206.5	17.8	205	22	AAE81105	Human JNK activati
31	205	17.7	184	21	AAE68795	Amino acid sequenc
32	205	17.7	184	22	AAE73216	Human phosphatase
33	205	17.7	184	22	AAE67167	Human dual-specifi
34	205	17.7	184	22	AAE66431	Human DSP-3 protei
35	200	17.2	184	22	AAE73213	Murine phosphatase
36	200	17.2	205	22	AAE66443	Murine DSP-3 varia
37	197	17.0	74	22	AAE04854	Human SGP014 phosph
38	195.5	16.8	179	22	AAE25664	Human protein sequ
39	195	16.8	80	22	AAE41142	Human polypeptide
40	192.5	16.6	169	22	AAE06783	Human dual-specifi
41	192.5	16.6	169	22	AAE66439	Human MAP-kinase p
42	192.5	16.6	302	22	AAE76876	Human lung tumour
43	192.5	16.6	392	18	AAE09039	Mitogen-activated
44	192.5	16.6	394	15	AAE63601	MAP-kinase-phospha
45	192.5	16.6	394	16	AAE78635	Partial MAP kinase

ALIGNMENTS

RESULT 1

AAE04837

ID AAE04837 standard; Protein: 220 AA.

XX AC AAE04837;

XX DT 10-SEP-2001 (first entry)

XX XX Human SGP003 phosphatase polypeptide.

XX DE Human: SGP003 phosphatase polypeptide; phosphatase-related disease;

XX KW immune-related disorder; ocular disease; organ transplant rejection;

XX KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;

XX KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;

XX KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;

XX KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;

XX KW attention disorder; cognition disorder; psychotic disorder; cytostatic;

XX KW neurological disorder; viricide; nootropic; cerebroprotective; therapy;

XX KW neuroprotective; antibacterial; vulnery; tranquiliser; antisthmatic;

XX KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;

XX KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;

XX KW MKP; migraine; chromosome CHR10.

XX OS Homo sapiens.

XX XX

XX FH Key Location/Qualifiers

XX FT Domain 1..173

XX FT Domain /label= Catalytic_domain

XX FT Domain 54..199

XX FT Domain /label= Phosphatase_domain

XX FT Misc-difference 85

XX FT /note= "Encoded by CTG"

XX XX

XX PN WO200146394-A2.

XX PD 28-JUN-2001.
 XX PF 21-DEC-2000; 2000WO-US34736.
 XX PR 21-DEC-1999; 99US-0173255.
 XX PR 28-DEC-1999; 99US-0175766.
 XX PR 25-JAN-2000; 2000US-0178078.
 XX PR 31-JAN-2000; 2000US-0179301.
 XX PA (SUGE-) SUGEN INC.
 XX PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;
 XX PI Flanagan P;
 DR WPI: 2001-418058/44.
 DR N-PSDB; AAD09495.
 XX Novel phosphatase polypeptide useful for treating cancers,
 PT immune-related diseases and disorders, cardiovascular disease, brain or
 PT neuronal-associated diseases and metabolic disorders
 XX Claim 7; Fig 2; 186pp; English.
 XX The present invention relates to phosphatase polypeptides, nucleotide
 CC sequences encoding them, as well as various products and methods useful
 CC for the diagnosis and treatment of various phosphatase-related diseases
 CC and conditions. Substance that modulates the activity of phosphatase
 CC polypeptide is used to treat immune-related diseases and disorders,
 CC cardiovascular disease, brain or neuronal-associated diseases and
 CC metabolic disorders, including cancers of tissues, cancers of
 CC haematopoietic origin, diseases of central and peripheral nervous
 CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
 CC amyotrophic lateral sclerosis, viral infections, infections caused by
 CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
 CC sexual dysfunction, mood disorders, attention disorders, cognition
 CC disorders, hypotension, hypertension, psychotic disorders, neurological
 CC disorders, dyskinesias and organ transplant rejection. The present
 CC amino acid sequence is human SGP003 phosphatase polypeptide. This
 CC sequence is classified as dual specificity phosphatase (DSP) and MAP
 CC kinase phosphatase (MKP). SGP003 gene maps to chromosomal position
 CC CHR10.
 XX Sequence 220 AA;
 SQ

Query Match 99.8%; Score 1159; DB 22; Length 220;
 Best Local Similarity 99.5%; Pred. No. 1.2e-127;
 Matches 219; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTSGEVKTSLKNAYSSAKRLSPKMEEGEEDYCTPGAFELERLFWKSGPOYTHVNEVWP 60
 DB 1 mtsgevktslknayssakrlspkmeegeedyctpgafelerlfwksgpythvnewwp 60
 QY 61 KLYIGDEATLDRYRLQKAGFTHYVNAAGRWNVDPDYPYRDMDIQYHGVEADLPTFD 120
 DB 61 klyigdeataldryrlqkagftvnaahgrwnvdpdyrdmdiqyhgveadlptfd 120
 QY 121 LSVFFYPAAAFIDRLSDHSHKILVHCVMGSRSRATLVLMIHKMTLVDAIQOVAKN 180
 DB 121 lsvffypaaafidrlsdhshkylvhcvmggrsratlvlmihkmtlvdaiqqvakn 180
 QY 181 RCVLPNRGFLKQRLDKOLVQRRRSQRQGEEDGREL 220
 DB 181 rcvlpnrgflkqrlkdqlvqrrrsqrqgeeedgrel 220

RESULT 2
 ID AAB18667
 XX AAB18667 standard; Protein; 221 AA.
 AC AAB18667;
 XX

DT 22-JAN-2001 (first entry)
 XX A human regulator of intracellular phosphorylation.
 DE Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;
 XX neurological disorder; Parkinson's disease; demyelinating disease;
 KW meningitis; developmental disorder; neuromuscular disorder; cancer;
 KW myasthenia gravis; cell proliferative disorder; actinic keratosis;
 KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;
 KW autoimmune disorder; inflammatory disorder; Addison's disease;
 KW acquired immunodeficiency disease; allergy; diabetes mellitus;
 KW rheumatoid arthritis; microbial infection; trauma.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 2
 FT /note= "potential phosphorylation site"
 FT Modified-site 9
 FT /note= "potential phosphorylation site"
 FT Modified-site 16
 FT /note= "potential phosphorylation site"
 FT Modified-site 21
 FT /note= "potential phosphorylation site"
 FT Modified-site 69
 FT /note= "potential phosphorylation site"
 FT Modified-site 170
 FT /note= "potential phosphorylation site"
 FT Modified-site 208
 FT /note= "potential phosphorylation site"
 FT Active-site 146..158
 FT /note= "tyrosine specific protein phosphatase active site"
 FT WO200055332-A2.
 XX 21-SEP-2000.
 XX 17-MAR-2000; 2000WO-US07277.
 XX 18-MAR-1999; 99US-0125593.
 XX 20-MAY-1999; 99US-0135049.
 XX 09-JUL-1999; 99US-0143188.
 XX (INCY-) INCYTE PHARM INC.
 XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;
 PI Lu DAM, Au-Young J;
 XX WPI: 2000-602121/57.
 DR N-PSDB; AAA75684.
 XX Novel human intracellular phosphorylation regulator polypeptides and
 PT polynucleotides for diagnosis, prevention and treatment of
 PT neurological, cell proliferative and autoimmune/inflammatory disorders
 XX Claim 1; Page 84-85; 96pp; English.
 PS The present sequence represents a human regulator of intracellular
 XX phosphorylation (HRIP). HRIP is useful for screening agonists and
 CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
 CC are useful for treating a disease or condition associated with
 CC decreased or increased expression of functional HRIP. Diseases treated
 CC or diagnosed include neurological disorders such as stroke, Parkinson's
 CC disease, demyelinating diseases, bacterial and viral meningitis and
 CC other developmental disorders of the central nervous system,
 CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
 CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
 CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
 CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
 CC inflammatory disorder such as Addison's disease, acquired
 CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,

CC rheumatoid arthritis, microbial infection and trauma.

XX Sequence 221 AA;

Query Match 88.4%; Score 1026.5; DB 21; Length 221;

Best Local Similarity 91.1%; Pred. No. 4.4e-112;

Matches 204; Conservative 1; Mismatches 12; Indels 7; Gaps 2;

QY 1 MTSEVKTSLKNAYSSAKRLSPKMEEGEEDYCTPGAFELERLFWKSGSPQYTHVNEVWP 60

Db 1 mtsevktslknayssakrlspkmeegeedyctpgafelerlfrwksgspqythvnevwp 60

QY 61 KLYIGDEATADLRYLQKAGFTVNLAAHGRWNVDTPGYVRMDIQYHGVAEADLPT-- 118

Db 61 klyigdeataldrylqkagftvlnaahgrwnvdtg---rllprghgvprrrgggttc 117

QY 119 --FDLSVFFYPAAAFIDRALSDHSHKILVHCVMGRSRSATLVLAYLMHKDMTLVDAIQ 176

Db 118 pfdlsvffypaaafidralsdhshkylvhcvmgrsrsatlvlaylmhkdmtlvdaiq 177

QY 177 VAKNRCVLPNRGELKQLRELDKQLVQORRRSQRDGEEEDGREL 220

Db 178 vaknrcvlpnrgrfklqreldkqlvqorrrsqrdgdeeddrel 221

RESULT 3

AAB73219

ID AAB73219 standard; Protein; 198 AA.

XX AAB73219;

DT 11-MAY-2001 (first entry)

DE Human phosphatase LOC51207_h.

XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
KW schizophrenia; hamartoma.

XX Homo sapiens.

XX W0200112819-A2.

XX 22-FEB-2001.

XX 11-AUG-2000; 2000WO-US22158.

XX 13-AUG-1999; 99US-0149005.

XX (SUGE-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;

XX WPI; 2001-211226/21.

XX N-PSDB; AAF63571.

XX New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizophrenia, cardiac
PT dysfunction and/or vascular disorders

XX Claim 6; Fig 5; 138pp; English.

XX The present invention relates to phosphatase proteins and coding
CC sequences. The present sequence is one such phosphatase. Phosphatases are
CC enzymes that catalyse the dephosphorylation of proteins modified by
CC phosphorylation of serine, threonine or tyrosine residues. The
CC phosphatases are useful for treating a variety of diseases: for example
CC cancer e.g. breast, uterine, prostate, head, neck, lung cancers,
CC synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,
CC hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,

CC glioblastoma, colorectal cancer and thyroid cancer, pathophysiological
CC hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
CC congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,
CC ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
CC Zonana syndrome, schizophrenia and hamartomas.

XX Sequence 198 AA;

Query Match 40.1%; Score 465.5; DB 22; Length 198;

Best Local Similarity 48.9%; Pred. No. 2.6e-46;

Matches 91; Conservative 36; Mismatches 58; Indels 1; Gaps 1;

QY 22 PKMEEGEEDYCTPGAFELERLFW-KGSPQYTHVNEVWPPLYIGDEATADLRYLQKAG 80

Db 12 pkmgavqaspyqpttlasiqlrlwvraqatlnhldewvpslfigdayaardksllqig 71

QY 81 FTHVLNAAHGRWNVDTPGYVRMDIQYHGVAEADLPTFDLSVFFYPAAAFIDRALSDH 140

Db 72 ithvnaaagkfvdgtgkfyrgmsleyygieddnpfddlsyflpvarylraalsvpq 131

QY 141 SKILVHCVMGRSRSATLVLAYLMHKDMTLVDAIQOAVAKNRCVLPNRGELKQLRELDKQL 200

Db 132 grvlvhcamgvsrsatlvlafimiyenmtlveaiqtvgahrnncpnsqflrqlgldhrl 191

QY 201 VQORRR 206

Db 192 gretgr 197

RESULT 4

AAY68779

ID AAY68779 standard; Protein; 248 AA.

XX AAY68779;

XX 16-MAY-2000 (first entry)

XX Amino acid sequence of a human phosphorylation effector PHSP-11.

XX Human; phosphorylation effector; PHSP; proliferative disorder;

XX immune disorder; neuronal disorder.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 5 /note= "potential phosphorylation site"

FT Modified-site 20 /note= "potential phosphorylation site"

FT Modified-site 36 /note= "potential phosphorylation site"

FT Active-site 166..220 /note= "tyrosine specific phosphatase active site"

FT Modified-site 208 /note= "potential glycosylation site"

FT Modified-site 210 /note= "potential phosphorylation site"

FT Modified-site 245 /note= "potential phosphorylation site"

FT Modified-site 245 /note= "potential phosphorylation site"

XX W0200006728-A2.

XX 10-FEB-2000.

XX 28-JUL-1999; 99WO-US17132.

XX 28-JUL-1998; 98US-0123494.

XX 14-SEP-1998; 98US-0152814.

XX 14-OCT-1998; 98US-0173482.

XX 03-NOV-1998; 98US-0106889.

XX 19-NOV-1998; 98US-0109093.

XX 22-DEC-1998; 98US-0113796.

Db 312 gretgr 317

RESULT 6

AA92177

ID AAY92177 standard; Protein; 221 AA.

XX AC

XX AAY92177;

XX 01-AUG-2000 (first entry)

XX Human cardiovascular system associated protein tyrosine phosphatase 4.

XX Cardiovascular system associated protein tyrosine phosphatase 4;

XX CSAPTP-4; cytosolic; immunomodulatory; antidiabetic; virucide;

XX hypotensive; cardiant; tyrosine phosphatase modulator.

XX Homo sapiens.

OS

XX Key

XX Domain 1..158

XX Location/Qualifiers

XX /label= "N-terminal_domain"

XX /note= "unique"

XX Modified-site 39..44

XX /note= "N-myristoylation site"

XX Domain 68..221

XX /note= "phosphatase catalytic active domain"

XX Modified-site 164..169

XX /note= "N-myristoylation site"

XX Modified-site 181..184

XX /note= "N-glycosylation site"

XX Modified-site 183..186

XX /note= "Casein kinase II phosphorylation site"

XX Modified-site 218..220

XX /note= "Protein Kinase C phosphorylation site"

XX WO200018990-A2.

XX 06-APR-2000.

XX 30-SEP-1999; 99WO-US22924.

XX 30-SEP-1998; 98US-0163833.

XX 30-SEP-1998; 98US-0164193.

XX (MILL-) MILLENNIUM PHARM INC.

XX Acton S;

XX WPI: 2000-293136/25.

XX N-PSDB: AAA09031, AAA09032.

XX New cardiovascular system associated protein tyrosine phosphatases are

XX used in the diagnosis and treatment of e.g. immune disorders,

XX anti-proliferative disorders, metabolic disorders and cardiovascular

XX disorders

XX Claim 8; Page 154-155; 156pp; English.

XX The CSAPTP nucleic acid and protein molecules are used to modulate

XX regulation of cellular processes. CSAPTP nucleic acid and protein

XX molecules and modulators of CSAPTP activity and expression can be

XX used to treat a subject with a disorder characterized by aberrant CSAPTP

XX expression or activity. These disorders can include an immune disorder,

XX an anti-proliferative disorder, a proliferative disorder e.g. renal and

XX lung carcinomas, a metabolic disorder e.g. diabetes, viral pathogenesis,

XX a neural disorder, a cardiovascular disorder e.g. hypertension or

XX coronary heart disease or a disorder rising from improper phosphorylation

XX of a phosphorylated protein. Anti-CSAPTP antibodies are used to isolate

XX CSAPTP by standard techniques, to facilitate the purification of natural

XX and recombinantly produced CSAPTP from cells, to detect CSAPTP protein in

XX cell supernatant for evaluating the level of CSAPTP expression and to

XX monitor protein levels in tissue as part of a clinical testing procedure

CC to determine efficacy of a treatment procedure. CSAPTP nucleic acids,

CC antibodies and protein molecules can be used in screening assays and in

CC predictive medicine e.g. prognostic assays, monitoring clinical trials

CC and pharmacogenetics.

XX SQ Sequence 221 AA;

Query Match 39.9%; Score 463.5; DB 21; Length 221;

Best Local Similarity 50.0%; Pred. No. 5.3e-46;

Matches 90; Conservative 34; Mismatches 55; Indels 1; Gaps 1;

Qy 22 PMKEEGEEEDYCTGAFELERLFW-KGSPQVTHYNEVMPKLYIGDEATFALDRYLQKAG 80

Db 35 pkihgavqaspyqpptlasqlrlwvraqatlnhldewpsiflgldayaardkskllqig 94

Qy 81 FTHVLNAAHGRNWNVDTPDYYRDMIOYHGVADDLPTFDLSVFFYPAAAFIDRLSDDH 140

Db 95 ithvvnaagkfvdtagkfyrgmsleyvgieadnppfdlsyflpvaryiraalsvpq 154

Qy 141 SKILVHCVMGRSRSATLVLAHMLHMKMTLVDAIQOVAKNRCVLPNRFGLKOLRELDKOL 200

Db 155 grvlvncamgvsrsatlvlafimiyenmtlvealqtqahrnpcnsgflrqlgidnrl 214

RESULT 7

AA73214

ID AAB73214 standard; Protein; 198 AA.

XX AC

XX AAB73214;

XX 11-MAY-2001 (first entry)

XX Murine phosphatase AA144705_m.

XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;

XX cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;

XX congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;

XX Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;

XX schizophrrenia; hamartoma.

XX Mus sp.

XX WO200112819-A2.

XX 22-FEB-2001.

XX 11-AUG-2000; 2000WO-US22158.

XX 13-AUG-1999; 99US-0149005.

XX (SUGE-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;

XX WPI: 2001-211226/21.

XX N-PSDB: AAF63566.

XX New protein phosphatase polypeptide for diagnosing and treating

XX phosphatase related disorders such as cancer, schizophrenia, cardiac

XX dysfunction and/or vascular disorders

XX Claim 6; Fig 5; 138pp; English.

XX The present invention relates to phosphatase proteins and coding

XX sequences. The present sequence is one such phosphatase. Phosphatases are

XX enzymes that catalyse the dephosphorylation of proteins modified by

XX phosphorylation of serine, threonine or tyrosine residues. The

XX phosphatases are useful for treating a variety of diseases: for example

XX cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,

XX synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,

XX hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,

XX glioblastoma, colorectal cancer and thyroid cancer, pathophysiological

CC hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
CC congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,
CC ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
CC Zonana syndrome, schizophrenia and hamartomas.
XX
SQ Sequence 198 AA;

Query Match 39.7%; Score 461; DB 22; Length 198;
Best Local Similarity 49.7%; Pred. No. 8.9e-46;
Matches 94; Conservative 32; Mismatches 59; Indels 4; Gaps 2;

Qy 22 PKMEEGEEEDYCTPGAFELERLFW-KGSPQYTHVNEVMPKLYIGDEATALDRYLQKAG 80
Db 12 pkihgvavspqptlasqlrlwrrtaatlhinevwnplfgdayaardkskllqig 71
Qy 81 FTHVLNAAHGRWNVDTPDYRDMDIQYHGVEADDLPTFDLSVFFYPAAAFIDRALSDDH 140
Db 72 ithvvnvaagkfvdgtgkfyrgtpleyygieadonpfldlsvhflpvaryirdalnipr 131
Qy 141 SKILVHCVNGRSRATLVLAIVLMHDKMTLVDAIQVAKNRCVLPNRCFLKQLRELDKOL 200
Db 132 srviivhcamgvsrsatlvlaflmifemntlydaigtqahrdicpnsgflrqlgldnrl 191

Qy 201 VQRRRSQR 209
Db 192 ---rretgr 197

RESULT 8
AAM39977
ID AAM39977 standard; Protein: 327 AA.

XX AAM39977;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3122.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

XX N-PSDB; AAI59133.

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 3122; 10078pp; English.
XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX Sequence 327 AA;

Query Match 39.7%; Score 460.5; DB 22; Length 327;
Best Local Similarity 48.9%; Pred. No. 2.1e-45;
Matches 91; Conservative 35; Mismatches 59; Indels 1; Gaps 1;

Qy 22 PKMEEGEEEDYCTPGAFELERLFW-KGSPQYTHVNEVMPKLYIGDEATALDRYLQKAG 80
Db 141 pkihgvavspqptlasqlrlwrrtaatlhinevwnplfgdayaardkskllqig 200
Qy 81 FTHVLNAAHGRWNVDTPDYRDMDIQYHGVEADDLPTFDLSVFFYPAAAFIDRALSDDH 140
Db 201 ithvvnvaagkfvdgtgkfyrgmsleyvgieadonpfldlsvflpvaryiraalsvpq 260
Qy 141 SKILVHCVNGRSRATLVLAIVLMHDKMTLVDAIQVAKNRCVLPNRCFLKQLRELDKOL 200
Db 261 grvlvhamgvsrsatlvlaflmiceenmlvealqtqahrdicpnsgflrqlgldnrl 320
Qy 201 VQRRR 206
Db 321 gretgr 326

RESULT 9

AAB28519
ID AAB28519 standard; Protein: 291 AA.

XX AAB28519;

XX 05-FEB-2001 (first entry)

XX Human DSP-7.

XX Mitogen-activated protein kinase; MAP kinase;
XX dual-specificity phosphatase; DSP-7; cytostatic; immunosuppressive;
XX antiallergic; apoptosis modulation; gene expression modulation;
XX Duchenne muscular dystrophy; cancer; graft-versus-host disease;
XX autoimmune diseases; allergy; metabolic disease.

XX Homo sapiens.

XX WO2000060098-A1.

XX 12-OCT-2000.

XX 07-APR-2000; 2000WO-US09257.

XX 07-APR-1999; 99US-0128207.

XX 25-MAY-1999; 99US-0135757.

XX (CEPT-) CEPTYR INC.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2002, 17:30:50 ; Search time 1322.25 Seconds
(without alignments)
14535.235 Million cell updates/sec

Title: US-09-847-519A-1

Perfect score: 1165

Sequence: 1 ggcagtggtggtggtggg.....aaaaaaaaaaaaaaaaaaaa 1165

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_cm:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htgo_hum:*
- 31: em_htgo_inv:*
- 32: em_htgo_rod:*
- 33: em_htg_hum:*
- 34: em_htg_inv:*
- 35: em_htg_rod:*
- 36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1096.6	94.1	1262	6	AX180878	AX180878 Sequence
2	589.4	50.6	159517	9	AC018511	AC018511 Homo sapi
3	208	17.9	1726	9	AY040091	AY040091 Homo sapi
4	204.2	17.5	636	6	AX180880	AX180880 Sequence
5	204.2	17.5	665	9	BC001613	BC001613 Homo sapi
6	201	17.3	1917	6	AX180879	AX180879 Sequence
7	199.4	17.1	904	6	AX086020	AX086020 Sequence
8	197.8	17.0	937	9	AB027004	AB027004 Homo sapi
9	197.8	17.0	1053	9	BC009778	BC009778 Homo sapi
10	193.8	16.6	1243	9	BC003115	BC003115 Homo sapi
11	175	15.0	1691	6	AX113043	AX113043 Sequence
12	165.2	14.2	597	6	AX086010	AX086010 Sequence
13	165.2	14.2	686	10	AB027003	AB027003 Mus muscu
14	165.2	14.2	1094	10	AF237620	AF237620 Mus muscu
15	145.8	12.5	4360	6	AX180877	AX180877 Sequence
16	133	11.4	861	6	U19731	U19731 Sequence 1
17	133	11.4	861	6	U19731	U19731 Sequence 1
18	133	11.4	1930	9	BC002682	BC002682 Homo sapi
19	129.4	11.1	1080	10	AF280809	AF280809 Mus muscu
20	119.4	10.2	240	6	AX180923	AX180923 Sequence
21	103.4	8.9	177597	2	AC013603	AC013603 Homo sapi
22	103.4	8.9	184022	2	AC015779	AC015779 Homo sapi
23	98	8.4	205347	2	AL392111	AL392111 Homo sapi
24	97.8	8.4	148849	9	AL158837	AL158837 Human DNA
25	96.2	8.3	165245	2	AL451050	AL451050 Homo sapi
26	78.6	6.7	4162	2	AC012984	AC012984 Drosophil
27	78.6	6.7	165097	3	AC018490	AC018490 Drosophil
28	78.6	6.7	298705	3	AE003510	AE003510 Drosophil
29	73.4	6.3	1467	5	AF167296	AF167296 Gallus ga
30	70.6	6.1	1236	9	S80632	S80632 threonine,
31	70.6	6.1	1238	6	AR075981	AR075981 Sequence
32	70.6	6.1	1238	6	AX063176	AX063176 Sequence
33	70.6	6.1	1849	9	BC002671	BC002671 Homo sapi
34	70.6	6.1	2064	6	AX063198	AX063198 Sequence
35	70.6	6.1	2109	6	AX063199	AX063199 Sequence
36	70.6	6.1	2234	9	HSU21108	U211108 Human dual
37	70.6	6.1	2234	11	G28599	G28599 human SFS S
38	70.6	6.1	2981	9	HSM802175	AL137704 Homo sapi
39	69	5.9	1619	6	AX063174	AX063174 Sequence
40	69	5.9	2241	9	HSU48807	U48807 Human MAP k
41	69	5.9	4637	6	AX063177	AX063177 Sequence
42	67.4	5.8	2436	10	AF013144	AF013144 Rattus no
43	66.4	5.7	94752	9	AC003098	AC003098 Homo sapi
44	65.8	5.6	69061	2	AC012986	AC012986 Drosophil
45	65.8	5.6	168469	3	AC007886	AC007886 Drosophil

ALIGNMENTS

RESULT 1

AX180878	LOCUS	AX180878	1362 bp	DNA	PAT	06-AUG-2001
DEFINITION	Sequence 5 from Patent WO0146394.					
ACCESSION	AX180878					
VERSION	AX180878.1	GI:15132706				
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
TITLE	1 (bases 1 to 1262)					
JOURNAL	Plowman,G.D., Martinez,R., Whyte,D., Manning,G., Sudarsanam,S., Hill,R.J. and Flanagan,P.					
FEATURES	Mammalian protein phosphatases					
source	Patent: WO 0146394-A 5 28-JUN-2001;					
	Sugen, Inc. (US)					
	Location/Qualifiers					
	1..1262					
	/organism="Homo sapiens"					

Qy 661 gctggtcctcccgaaaccggggcttttgaagcagctccgggagctggacaagcagctgg 720


```
FEATURES             Location/Qualifiers
  source              1..636
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
BASE COUNT          123 a   207 c   182 g   124 t
ORIGIN

Query Match          17.5%; Score 204.2; DB 6; Length 636;
Best Local Similarity 64.8%; Pred. No. 2.1e-28;
Matches 321; Conservative 0; Mismatches 168; Indels 6; Gaps 1;

Qy 231 gcctttgagctggagcgctcttctggaaggcgagctccagtagcacccagctcaacgag 290
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 GTCCTGAGTTGGAGCGGCTCTCTACACAGGCAAGACAGCGCTGTAACTATCCGCGACGAG 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 291 gtcctggcccaagctctacattggcgatgagcgagcgctggagccgctatagctcgag 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 GTCGTGGCCAGGCTCTATCTCGGACACGAGCATGGCTAAACAACCGCGGAGCTTCGC 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 351 aaggcggttcacgacgtgctgaagcggtgacggcccgctggaacgtggacactgag 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 253 CGCTGGGCAATCAGCAGCTCTCAATGCTTCACACAGCGGCTGGCGAG-----GCACG 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 411 cccgactactaccgacatggacatcagtagcagcgctgagcgagcgagcagcctgccc 470
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 CCCGAGGCTATAGGGGCTGGCATCCGCTACCTGGGTGTGAGGCCACGACTCGCCA 366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 471 accttgacctagtgctttcttaccggcgagcagccttcctcatcgacagcgctaaagc 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 367 GCCTTTGACATGAGCATCCACTTCACAGCGCTGCCGACTTCATCCACCGGCGCTGAGC 426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 531 gacgaccacagtagatcctgttctactgctcagtcagtcagtcagtcagtcagtcagtc 590
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 427 CAGCAGGAGGAGATCCTGGTGCATGTCTGTGGCGTGTGAGCGGATCCGCCACCCCTG 486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 591 gtcctggcctacctgatgatacacaagacatgacacctggtggagcgcctaccagcaagt 650
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 487 GTACTGGCTACCTACCTGTACACACACCTTACCTCTGTGGAGGCCATCAAGAAAGTC 546
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 651 gccaaagaacctgctgctctcccaacggggcttttgaagcgagcctccggagctggac 710
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 547 AAAGACCACGAGGATCATCCCAACGGGGCTTCTGTAGGCGAGCTCTGGCCCTGGAC 606
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 711 aagcagctggtgag 725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 607 CGCAGGCTGGCGAG 621
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
BC001613 1665 bp mRNA PRT 12-JUL-2001
LOCUS Homo sapiens, clone MGC:1136 IMAGE:3535215, mRNA, complete cds.
ACCESSION BC001613
VERSION BC001613.1 GI:12804418
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1665)
Direct Submission
Strausberg, R.
Submitted (03-JAN-2001), National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: qcaps-r@mail.nih.gov
Tissue Procurement: DCTD/dnp
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```


Db 880 AAGACACCGAGGAGCATATCCCAACCGGGGCTTCTGAGGAGAGCTCTCTGGCCCTGGAC 939
QY 711 aagcagctgtgtcag 725
Db 940 CGCAGGCTGGCGCAG 954

RESULT 6
LOCUS AX180879 1917 bp DNA PAT 06-AUG-2001
DEFINITION Sequence 6 from Patent WO0146394.
ACCESSION AX180879
VERSION AX180879.1 GI:15132707
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE Plozman,G.D., Martinez,R., Whyte,D., Manning,G., Sudarsanam,S.,
Hill,R.J. and Flanagan,P.
JOURNAL Mammalian protein phosphatases
Patent: WO 0146394-A 6 28-JUN-2001;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
source 1..1917
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 368 a 596 c 565 g 388 t
ORIGIN

Query Match 17.3%; Score 201.; DB 6; Length 1917;
Best Local Similarity 63.0%; Pred. No. 6.9e-28;
Matches 328; Conservative 0; Mismatches 190; Indels 3; Gaps 1;

QY 192 gaggaaggaggagaggagactactgcacccctggagcctttgagctggagcgctc 251
Db 52 GAGCTGGGGGAGAGGACAAAGCCACGCTTGGCCCGAGCATCTGGAGCTGGAGGAGCTC 111
QY 252 ttctgaaggagctccccagtcacacccagtcacacagagcttgccccaaagctctacatt 311
Db 112 CTCGGGGCAGGGAAGTCTTCTTCACGCGCTGTGGAGCAAGTTTGGCCCAACCTTTTCATA 171
QY 312 ggcgatgagcgagcgctggaccgctataggctcagaagcggggttcacgcagctg 371
Db 172 GGAGATGGGGCCACGGCAACACCGCTTTACGCTGTGGAGGTGGGCATCACCCACGTG 231
QY 372 ctgaacgcggccacggcgcgtggaacgtggacactgggcccgaactactaccgcgacatg 431
Db 232 CTGAACGCGCCCAAGGGCCCTTACTGTACGGCGGCCCTGACTTCTACGGCAGCAGT 291
QY 432 gacatcagtcacacggctggagggcgacagacctcccacaccttcagctcagtgcttc 491
Db 292 GTGAGC---TACCTGGGGGTGGCAGGCCACGACCTCCCTGATTTTGACATCATGTGCTAC 348
QY 492 ttctaccgcgcgacgcttcacagacgagcgttaagcagacacacagtagatcctg 551
Db 349 TTCTCTCTCGCGCTGACTTCATCCACCGTGGCCCTCAACACGCGTGGGGCCAAAGGCTCTG 408
QY 552 gttcactcgtgtcagggcgagcgctgagccacacctggtcctggcctacacctgatgc 611
Db 409 GTGCACTGTGTGGCGGTGAGCCGCTCTGCCACGCTGTGCTCTGGCTTACCTCATGCTG 468
QY 612 cacaagagatgacctgtgtgagcgccatccacagagtggtgcccagaaacgcgtcgtcctc 671
Db 469 CACCAGCGGCTGTCTCTGGCCAGGGGGTGTATCACCCTGTAGGAGCAGCAGCATGGGTCTTC 528
QY 672 ccgaacggggcttttgaagcagctccggagagctgagacaa 712
Db 529 CCCAACCGAGGCTTCTCTGCACACGAGCTCTGCAGGCTGGACCA 569

RESULT 7
LOCUS AX086020 904 bp DNA PAT 09-MAR-2001
DEFINITION Sequence 17 from Patent WO0112819.
ACCESSION AX086020
VERSION AX086020.1 GI:13275841
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE Plozman,G.D., Martinez,R., Whyte,D., Hill,R., Flanagan,P. and
Lioubin,M.
JOURNAL Protein phosphatases and diagnosis and treatment of
phosphatase-related disorders
Patent: WO 0112819-A 17 22-FEB-2001;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
source 1..904
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 183 a 282 c 249 g 190 t
ORIGIN

Query Match 17.1%; Score 199.4; DB 6; Length 904;
Best Local Similarity 64.3%; Pred. No. 1.6e-27;
Matches 299; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 276 acccagctcaacgaggtctgtgcccagctctacatctgagcgagcgcgctggac 335
Db 200 AACCATATCGATGAGGTCTGCCCCAGCCTCTTCTGGGAGATGCGTACGAGCCCGGAC 259
QY 336 cgtataggctgcagaagcggggttcacgcagcgtgctgaacggcgccacggcgctgg 395
Db 260 AAGAGCAAGCTGATCCAGCTGGGAATCACCCAGCTTGTGAATGCCGTGCAGGCAAGTTC 319
QY 396 aagctgagacactggccgagactactaccgcagacatgacatccagtcacgagcgctgag 455
Db 320 CAGGTGGACACAGGTGCCAAATTTCTACCGTGGAAATGTCCTGGAGTACTATGGCATTTGAG 379
QY 456 gccagcagacctgcccacttgcacacctgagtgctcttctaccgcggcgagccttcac 515
Db 380 GCGAGCAGCAACCCCTTCTTCAGCTCAGTGCTACTTTCTGCTGTGTCGATACATC 439
QY 516 gacagagcgctagcgagcagcaccagtaagatcctggttcaactgcgtcatgpgccgcagc 575
Db 440 CGAGCTGCCCTCAGTGTTCGCCAAGCGCGGTGCTGTGTACACTGTGTCATGGGGTAAAG 499
QY 576 cggtcagcaccctgtgctgacctacgtgatccacaagacatgacacctggtggac 635
Db 500 CGCTCTGCCACACTGTGCTGGCCCTTCTCATGATCTATGAGAACATGACGCTGGTAGAG 559
QY 636 gccatccagcaagtggtgcccagaagcgtgctcctcccgacccgggctcttttgaagcag 695
Db 560 GCATCCAGCGGTGCGAGGCCACCGCAATATTCGCCCTAACTCAGGCTTCTCTCCGCGAG 619
QY 696 ctccggagctggaacagcagctggtgcagcagagcgagcggttc 740
Db 620 CTCAGGTTCTGGACACACCGACTGGGGCGGAGACGGGGCGGTTC 664

RESULT 8
LOCUS AB027004 937 bp mRNA PRI 13-JAN-2000
DEFINITION Homo sapiens mRNA for protein phosphatase, complete cds.
ACCESSION AB027004
VERSION AB027004.1 GI:6692781
KEYWORDS protein phosphatase.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Homo sapiens, protein phosphatase, clone MGC:13471 IMAGE:4047963,
mRNA, complete cds.
BC009778
BC009778.1 GI:14602534
MGC.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
Strausberg, R.
Direct Submission
Submitted (02-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.

CDNA LIBRARY Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbiology.org>

contact: amandajstembiroj.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Ketteman and Anuradha Madan

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA q: 6692781; clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Series: IRAL Plate: 19 Row: b Column: 2

passed the following selection criteria: matched mRNA q1: 6692781.

```

1. 1000
/organism="Homo sapiens"

```

```
/db_xref="LocusID:51207"  
/db_xref="taxon:9606"  
/db_xref="taxon:9606"
```

```
/cd=/usr/local/arc/arc3000
/clone="MGC:13471 IMAGE:4047963"
```

```
/tissue_type="Bone marrow, chronic myelogenous leukemia"  
/clone_lib="NTH MGC 54"
```

```
/lab_host="DH10B"
```

```

/NOTE= vector: PDNR-LTB
190. .786

```

```
/codon_start=1
/product="prot of n pbac-hat-c"
```

```

/Protein Phosphatase
/protein_id="AAH09778.1"

```

```

/db_xref="GI:14602535"
/translation="MPSLOKQNI.RRPKIHCAVQASPVQDPTIASIQRII WYPOAATIN

```

HIDEVWPSLEGLDAYAARDKSKLIQLGITHVVNAAAGKFQVDTGAKFYRGMSLEY YGI

EADDNPEFFDLSVYFLPVARY IRAALSVPQGRVLVHCAMGVRSATLVLAFLMIYENMT
LVEA IOTVQVAHRNICPNSGELBOLOVLDNRLGRETGRF"

·	248 a	298 c	298 g	209 t
---	-------	-------	-------	-------

C
L
C
r

T
v

C
P
P

C
M
C
r

F
I
E
O

G
C
P
r

J

1 Similarity 64.18; Pred. No. 3e-27;
17.08; Score 197.8; DB 9; Length 1033;

298; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

cccaagtcacgaggtctggcccaagctctacattggcgatgaggcagggcgctggac 335

[illegible]

.....

gctataggctgcagaaggcggggtcacgcacgtgctgaacgcggccacggccgctgg 395

AGAGCAAGCTGATCCAGCTGGGAATCACCCACGTTGTGAATGCCGCTGCAGGCAAGTTC 438

acgtggacactggggccqactactaccggacatggacatccagttaccacggcgtggag 455

[illegible][illegible]

ccgacgacctgcccacctcgacctcagtgctctcttctaccggcgagccttcac 515

Db 499 GCGGATGACACCCCTTCTTCGACCTCAGTGTCTACTTTCTGCTGTCTCGATACATC 558
Qy 516 gacagagcgttaagcagaccacagtaagatcctgttctactcgtcgtcagcgcgcagc 575
Db 559 CGAGCTCCCTCAGTGTTCGCCAAGCGCGTGTGTATACATCTGCCATGGGGTAAAGC 618
Qy 576 cggtcagccacctggtctcctggtcctactgatcaccagagacatgacctggtgagc 635
Db 619 CGCTCTGCCACACTGTCTTCGCGCTTCTCATGATCTATGAGACATGACGCTGTGAGAG 678
Qy 636 gccatccagcaagtgcccaagaccgctgctcctccgaaccgggctttttgaaagcag 695
Db 679 GCCATCAGAGCGGTGCGAGGCCACCGCAATATCTGCCCTAACTCAGGCTTCTCTCCGCGAG 738
Qy 696 ctccggagctggacaagcagctgctgagcagagcagcagcgtcc 740
Db 739 CTCCAGGTTCTGGACAACCGACTGGGCGGAGACGGGGCGGTTC 783

RESULT 10

BC003115 1243 bp mRNA PRI 12-JUL-2001
LOCUS Homo sapiens, clone MGC:2627 IMAGE:3502718, mRNA, complete cds.
DEFINITION BC003115
ACCESSION BC003115
VERSION BC003115.1 GI:13111894
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg, R.
Direct Submission
Submitted (13-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: d Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: GenomeScan gene
prediction.

FEATURES

source
1..1243
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:2627 IMAGE:3502718"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH_MGC_19"
/lab_host="DH10B-R"
/note="vector: pOTB7"
279..539
/codon_start=1

CDS

/product="Unknown (protein for MGC:2627)"
/protein_id="AAH03115.1"
/db_xref="GI:13111895"
/translation="MSIHROTADFTIHRALSQPGKILVHCIVGVSRSATLVLAYLML
YHHLTIVEAIKKVKDHRGIIPNRGFLQRLALDRURQGLEA"
BASE COUNT 298 a 346 c 379 g 220 t
ORIGIN
Query Match 16.6%; Score 193.8; DB 9; Length 1243;
Best Local Similarity 64.8%; Pred. No. 1.6e-26;
Matches 321; Conservative 0; Mismatches 167; Indels 7; Gaps 2;
Qy 231 gctttgagctggagcgcgtctctggaaggcagctcccccagtagcacacgctcaacag 290
Db 37 GTCTTCGAGTTGGAGCGGCTCTCTACACAGGCAAGACAGCCTGTAAACATGCCGACGAG 96
Qy 291 gtctggcccaagctctacattggcgatgagcgagcgctgagccgctatagcgtcag 350
Db 97 GTCTGGCCAGGCTCTTATCTCGGAGACACGAGACATGGCTAAACAACCGCGGGAGCTTCCC 156
Qy 351 aaggcggggttcacgacgctgctgaacgcgccacgcgctggaacgctggacactggg 410
Db 157 GCCTTGGCATACGACGCTCTCAATGCCCTCACACAGCCGGTGGCGAG-----GCACG 209
Qy 411 ccgactactaccgcgacatggacatccagtagccagcgtgagcgccgagcagcctgcc 470
Db 210 CCCGAGGCTATGAGGGGTGGGCTCCGCTACCTGCTGTGAGGGCCAGCATCGCCA 269
Qy 471 accttcgacctcagctctcttcttaccggcgccgagccttcacgacgagcgaagc 530
Db 270 GCCTTTGACATGACATCCACTTCCAGACGGCTGCCGACTTCATCCACCGGGCGCTGACG 329
Qy 531 gacgaccacagtagatcctcgtggttcactgctgctgagcgcgcgcgcgcgcgcctg 590
Db 330 CAGCCAGGAGGGAAGATCCTGTGCTGCTGTGGCGGTGAGCCGATCCGCCACCCCTG 389
Qy 591 gtctcgtcttaccgtgatcccaagagacatgacctggtgagcgccacgacgaagt 650
Db 390 GTACTGGCTACCTACATGCTGTACACACCTTACCTCTGTGGAGGCCATCAAGAAAGTC 449
Qy 651 gccaaagaccgctgctcctcccgaaacggggttttgaagcagcgtccggagcgtgagc 710
Db 450 AAAGACCCAGGAGCATATCCCAACCGGGGCTTCTGTAGGCGAGCTCTCTGGCCCTGGAC 509
Qy 711 aagcagctggtgcag 725
Db 510 CGCAGGCTGGCGGAG 524
RESULT 11
ARL13043
LOCUS ARL13043 1691 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 12 from patent US 6132964.
ACCESSION ARL13043
VERSION ARL13043.1 GI:14093365
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1691)
AUTHORS Bandman, O., Lal, P., Hillman, J.L., Corley, N.C., Guegler, K.J. and
Shah, P.
TITLE Hydrolase enzymes
JOURNAL Patent: US 6132964-A 12 17-OCT-2000;
FEATURES Location/Qualifiers
source
1..1691
/organism="unknown"
BASE COUNT 345 a 493 c 528 g 325 t
ORIGIN
Query Match 15.0%; Score 175; DB 6; Length 1691;

BASE COUNT 144 a 205 c 182 g 155 t
ORIGIN

Query Match 14.2%; Score 165.2; DB 10; Length 686;

Best Local Similarity 57.9%; Pred. No. 3.9e-21;

Matches 293; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

```
QY 231 gctttagctgagcgcctctctggaaggcagtcctccagtcacacccagtcacagcag 290
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 85 GCCTCTCTGAGCGATTGCTGTGGGTCGCGGACTGCCACATGACCCACATCAATGAG 144
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 291 gtctggcccaagctctacattggcgatgagcgagcgctgacccgctatagctgcag 350
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 145 GTCTGGCCCAACCTTTCTTGGGAGATGCGTATCTGCCAGAGACAAGGGTCTCTAATC 204
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 351 aagcggggttcacgcacgtgctgaacgcggccacgcggcgcgtggaacgtggacactggg 410
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 205 CAGCTGGGCATTACCCCATGTTGTAATGTGGCTGCGGGCAAGTTCCAGGTGGACACAGGT 264
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 411 ccgactactaccgcgacatggacatccagtcacacgcgctgagcgccgacacactccc 470
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 265 GCGAAGTTTACCGTGGAGAACACCTCTGGAGTACTATGGCATGAGGCTGATGACAACCCC 324
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 471 accttcacactcagtgctctcttctacccgcggcagcgccttcctcgcacagagcgctaacg 530
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 325 TTCCTTACCTCAGCGTCCACTTTCTGCTGTTGCTGTACATCAGAGATGCCCTCAAT 384
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 531 gacgacacagtaagaacctggttcactgcgtcatggccgcgcgcgcgcgcgcgcgcgcgc 590
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 385 ATTCCCGGAAGCGAGTGTGCTGCTACCTGCGCTATGGGGGTGAGTCTGTCGCACAAAT 444
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 591 gtctgcctaccctgatgatcccaagacatgacacccctggtgacgcacccatccagcaagt 650
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 445 GTCTTGGCCCTCCATCATGATCTTCGAGAATGATGACATGATGATGATGATGATGATGATG 504
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 651 gcaagaacgcgtgcgtctctccggaacgcggggtttttgagcagcgtccggagcgtggac 710
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 505 CAGGCCACCGAGATATCTGCCCACTCAGGCTTCTCCGACAGCTCCAGGTCTTGAC 564
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 711 aagcagctggtgcagcagagcgacg 736
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 565 AACAGGCTGAGGGCGGGAACAGGAAG 590
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 14

AF237620 1094 bp mRNA ROD 02-MAR-2001
LOCUS Mus musculus dual-specificity phosphatase TS-DSP6 mRNA, complete cds.

ACCESSION AF237620.1 GI:13183066

VERSION 2

KEYWORDS

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1094)

Aoyama, K., Matsuda, T. and Aoki, N.

Direct Submission

Submitted (19-FEB-2000) Applied Molecular Biosciences, Nagoya

University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan

Location/Qualifiers

1. .1094

/organism="Mus musculus"

/db_xref="taxon:10090"

203. .799

/codon_start=1

/product="dual-specificity phosphatase TS-DSP6"

CDS

203. .799

/codon_start=1

/product="dual-specificity phosphatase TS-DSP6"

/protein_id="AAK15037.1"

/db_xref="GI:13183067"

/translation="MDSLQKQLRRPKIHCAVOVSPQPTLASLQRLWVRRATLIT
HINEVMPNLELGDAYAARDKRLIQGITHVNVAAKQFOVDGAKFYRTPLEYGI
EADDNPFDFLSVHFLPVARYIRDALNIPRSVLVHCAMGVSRSATIVLAFIMFENMT
LVDAIQTVAHRDICPNSGFLRLQVLNRLRRETGL"

BASE COUNT 287 a 293 c 284 g 227 t 3 others

ORIGIN

Query Match 14.2%; Score 165.2; DB 10; Length 1094;
Best Local Similarity 57.9%; Pred. No. 3.5e-21;

Matches 293; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

```
QY 231 gctttagctgagcgcctctctggaaggcagtcctccagtcacacccagtcacagcag 290
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 287 GCCTCTCTGAGCGATTGCTGTGGGTCGCGGACTGCCACATGACCCACATCAATGAG 346
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 291 gtctggcccaagctctacattggcgatgagcgagcgctgacccgctatagctgcag 350
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 347 GTCTGGCCCAACCTTTCTTGGGAGATGCGTATCTGCCAGAGACAAGGGTCTGCTAATC 406
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 351 aagcggggttcacgcacgtgctgaacgcggccacgcggcgcgcgcgcgcgcgcgcgcgcgc 410
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 407 CAGCTGGGCATTACCCCATGTTGTAATGTGGCTGCGGGCAAGTTCCAGGTGGACACAGGT 466
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 411 ccgactactaccgcgacatggacatccagtcacacgcgctgagcgccgctatagctgcag 470
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 467 GCGAAGTTTACCGTGGAGAACACCTCTGGAGTACTATGGCATGAGGCTGATGACAACCCC 526
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 471 accttcacactcagtgctctcttctacccgcggcagcgccttcctcgcacagagcgctaacg 530
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 527 TTCCTTACCTCAGCGTCCACTTTCTGCTGTTGCTGTACATCAGAGATGCCCTCAAT 586
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 531 gacgacacagtaagaacctggttcactgcgtcatggccgcgcgcgcgcgcgcgcgcgcgcgc 590
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 587 ATTCCCGGAAGCGGAGTGTGCTGCTACCTGCGCTATGGGGGTGAGTCTGCCACAAAT 646
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 591 gtctgcctaccctgatgatcccaagacatgacacccctggtgagcgccatccagcaagt 650
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 647 GTCTTGGCCCTTCCATCATGATCTTCGAGAATGATGACATGATGATGATGATGATGATGATG 706
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 651 gcaagaacgcgtgcgtctctccggaacgcggggtttttgagcagcgtccggagcgtggac 710
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 707 CAGGCCACCGAGATATCTGCCCACTCAGGCTTCTCCGACAGCTCCAGGTCTTGAC 766
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 711 aagcagctggtgcagcagagcgacg 736
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 767 AACAGGCTGAGGGCGGGAACAGGAAG 792
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15

AX180877

LOCUS AX180877

DEFINITION Sequence 4 from Patent WO0146394.

ACCESSION AX180877

VERSION AX180877.1 GI:15132705

KEYWORDS human.

SOURCE Homo sapiens

REFERENCE 1 (bases 1 to 4360)

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

Plowman, G.D., Martinez, R., Whyte, D., Manning, G., Sudarsanam, S.,

Hill, R.J. and Flanagan, P.

Mammalian protein phosphatases

Patent: WO 0146394-A 4 28-JUN-2001;

Sugen, Inc. (US)

Location/Qualifiers

1. .4360

/organism="Homo sapiens"

/db_xref="taxon:9606"

1138 a. 1076 c 1363 g 783 t

BASE COUNT

	Query Match	12.5%	Score 145.8	DB 6	Length 4360
	Best Local Similarity	58.2%	Pred. No. 1.1e-17		
	Matches 296	Conservative	0	Mismatches 207	Indels 6
					Gaps 2
QY	280	acgtcaacgaggtctgcccgaagctctacatgctgatgagcgacgagcgctgaccgct	339		
DB	761	AGGTGGATGAGGTCTGCCCAATCTTCATAGCTGAGAAGAGTGTGGCTGTGAACAAGG	820		
QY	340	ataagctgcagaagcgggggtttcacgcacgtgctggaacgcgccacgcccgcctggaacg	399		
DB	821	GGAGGCTGAAGAGGCTGGGAATCACCCACATCTCTGAATGCTGCCATGGCACCGG	877		
QY	400	tggacatggggccgactactaccgcgacatggacatccagttaccacggcggtggagccg	459		
DB	878	TTTACACTGGCCCCGAATTTACACTGGCCTGGAGATCCAGTACCTGGGTAGAGGTGG	937		
QY	460	acgacctgccacacttcgacct---cagtgctctctctaccgcggcgagccttcctacg	516		
DB	938	ATGACTTTCTTGAGGTGGACATTTCCGACGATTTCCGGAGGCGTACTGTCTATTACATCA	997		
QY	517	acagagcgctaagcgcgacccacagtaagaacctgggttcactggtcgtcatggcgccgacgc	576		
DB	998	TTTTCTCTGTGTGTTTCATTTCAGGGAAAGTCTTGTCAGCAGCGGAAATGGCATCAGCC	1057		
QY	577	ggtcagccacctggtcctggcctacctgatgatccacaaggacatgacctggtgagcg	636		
DB	1058	GGTCAGCAGTGTGCTGTGCTGCGCTTACCTGTATGATCTTCCACAACATGCCATCTCTGGAGG	1117		
QY	637	ccatccagcaagtggccaagaacgcgtcgctctctccgaacccggggtttttggaacgc	696		
DB	1118	CTTTTGATGACCGTGGTAAAGAGCGGGCCATCTTACCCTCAATGACGGCTTCTCTGAAGCAGC	1177		
QY	697	tccgggagctggacaacgacgtggtgacgacagggcgacggtccacgcgcgcaggaagcggtg	756		
DB	1178	TGCGGGAGCTCAATGAGAAGTTGATGGAGGAGAGAGAAGAGACTATGCGCGGGAGGGGG	1237		
QY	757	aggaggaggtatggcagggagctgtaggcc	785		
DB	1238	GATCAGCTGAGGCTTGAGGAGGGCGGCGAGGC	1266		

Search completed: February 8, 2002, 18:42:47
Job time: 4317 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:40:01 ; Search time 12.62 Seconds
(without alignments)
392.292 Million cell updates/sec

Title: US-09-847-519A-2
 Perfect score: 1161
 Sequence: 1 MTSGEVKTSLNAYSSAKRL.....VQRRRSQRDGEEDGREL 220

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```
Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB pep.*
5: /cgn2_6/ptodata/2/iaa/PCUTS_COMB pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1 pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	465.5	40.1	198	4	US-09-163-833-2	Sequence 2, Appli
2	426	36.7	207	4	US-09-013-881-4	Sequence 4, Appli
3	363	31.3	263	4	US-09-164-193-5	Sequence 5, Appli
4	346.5	29.8	185	1	US-07-988-273-2	Sequence 2, Appli
5	346.5	29.8	185	3	US-07-848-810-25	Sequence 25, Appli
6	346.5	29.8	185	4	US-09-164-193-21	Sequence 21, Appli
7	346.5	29.8	185	5	PC7-US93-12019-2	Sequence 2, Appli
8	247.5	21.3	118	1	US-07-988-273-3	Sequence 3, Appli
9	247.5	21.3	118	5	PC7-US93-12019-3	Sequence 3, Appli
10	192.5	16.6	393	2	US-08-990-379-4	Sequence 4, Appli
11	192.5	16.6	394	2	US-08-530-290-23	Sequence 23, Appli
12	192.5	16.6	395	2	US-08-990-379-5	Sequence 5, Appli
13	188.5	16.2	313	2	US-08-990-379-7	Sequence 7, Appli
14	184.5	15.9	314	4	US-09-164-193-22	Sequence 22, Appli
15	177.5	15.3	198	4	US-09-045-973-5	Sequence 5, Appli
16	169	14.6	367	2	US-08-530-290-24	Sequence 24, Appli
17	168	14.5	226	4	US-09-045-973-8	Sequence 8, Appli
18	166	14.3	367	2	US-08-990-379-6	Sequence 6, Appli
19	157.5	13.6	397	2	US-08-990-379-8	Sequence 8, Appli
20	142.5	12.3	117	1	US-07-988-273-4	Sequence 4, Appli
21	142.5	12.3	117	5	PC7-US93-12019-4	Sequence 4, Appli
22	128.5	11.1	353	4	US-09-013-881-3	Sequence 3, Appli
23	123	10.6	72	2	US-08-530-290-20	Sequence 20, Appli
24	113	9.7	72	2	US-08-530-290-19	Sequence 19, Appli
25	102.5	8.8	201	2	US-08-933-750C-21	Sequence 21, Appli
26	102.5	8.8	201	4	US-09-234-613-21	Sequence 21, Appli
27	100.5	8.7	116	1	US-07-988-273-5	Sequence 5, Appli

28	100.5	8.7	116	5	PCT-US93-13019-5	Sequence 5, Appl
29	97.5	8.4	1442	1	US-08-015-386A-3	Sequence 3, Appl
30	97.5	8.4	1442	2	US-08-446-363-8	Sequence 3, Appl
31	93.5	8.1	245	1	US-08-015-985-8	Sequence 8, Appl
32	93.5	8.1	263	1	US-08-685-992-5	Sequence 5, Appl
33	93.5	8.1	263	2	US-09-144-925-5	Sequence 5, Appl
34	93.5	8.1	1445	1	US-08-015-986A-2	Sequence 2, Appl
35	93.5	8.1	1445	2	US-08-446-363-2	Sequence 2, Appl
36	85	7.3	173	3	US-08-725-532A-6	Sequence 6, Appl
37	85	7.3	173	4	US-09-164-193-20	Sequence 20, Appl
38	84	7.2	173	4	US-09-164-193-2	Sequence 2, Appl
39	83.5	7.2	150	4	US-09-164-193-23	Sequence 23, Appl
40	80.5	6.9	563	4	US-08-987-691A-2	Sequence 2, Appl
41	79.5	6.8	242	1	US-08-015-985-7	Sequence 7, Appl
42	79.5	6.8	360	2	US-08-531-927B-2	Sequence 2, Appl
43	79.5	6.8	360	4	US-09-041-686-13	Sequence 13, Appl
44	79.5	6.8	2308	1	US-08-015-973-1	Sequence 1, Appl
45	79.5	6.8	2308	2	US-08-448-164-1	Sequence 1, Appl

ALIGNMENTS

```

RESULT      1
US-09-163-833-2
/ Sequence 2, Application US/09163833
/ Patent NO. 6268135
/ GENERAL INFORMATION:
/ APPLICANT: Acton, Susan
/ TITLE OF INVENTION: NO. 6268135el
/ FILE REFERENCE: mni-059
/ CURRENT APPLICATION NUMBER: US/09/
/ CURRENT FILING DATE: 1998-09-30
/ NUMBER OF SEQ. ID NOS: 3
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 198
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-163-833-2

```

```

Query Match      40.1%; Score 465.5; DB 4; Length 198;
Best Local Similarity 48.9%; Pred. No. 2.8e-46;
Matches 91; Conservative 36; Mismatches 58; Indels 1; Gaps 1;

QY 22 PKMEEEGEEDYCTPGAFALERLFW-KGSPQYTHVNEVPKLYIGDEATALDRYRLOKAG 80
Db 12 PKTHGVAOASPYOPPTLASLQRLLLVWROATLNLHIDEVWPSLFGDAAARDKSKLLQLG 71

QY 81 FTHVLNAAGRWNVDPGYRRMDIQYHGVAEADDDPTFDLSVFFYPAAAFIDRALSDDH 140
Db 72 ITHVNAAGKFQVDTCAGEYKRGMSLEYEYIEADDNPFDFLSVYFLPVARYIRAALSPVQ 131

QY 141 SKLLVHCVMGRSRSATLVLLAYLATHKDMTLDALQOAVAKNRCVLPNRFGLKQLRELDKOL 200
Db 132 GRVLLHCAMGVSRSATLVLAFLMAYENMTLVYRAIQTVQAHNRNCPNSGFLRQLQVLDNRL 191

QY 201 VOORER 206
Db 192 GRETGR 197

RESULT 2
US-09-013-881-4
; Sequence 4, Application US/09013881
; Patent No. 6132964
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guebler, Karl J.

```

APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013,881
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LUCY J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0470 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: MUSCNOT02
CLONE: 971204
US-09-013-881-4

Query Match 36.7%; Score 426; DB 4; Length 207;
Best Local Similarity 53.0%; Pred. No. 1.2e-41;
Matches 87; Conservative 26; Mismatches 45; Indels 6; Gaps 2;

QY 39 FELERLFWKGSPOYTHVNEVWPVKLYIGDEATLDRLYRLOKAGFTHTLVNAHGRNVDGTP 98
Db 46 FELERLYTGTACNHADDEVWPGLYLGDQDMANNRRLRLGITHVLNASHSRWR--GTP 103
QY 99 DYIRMDIQYHGVEADDLPFDLSVFYPAAAFIDRLSDHSLVHCVMGSRSRATLV 158
Db 104 EAYEGLGTRYLGE----PAFDMSTHFQTAADFTHRLSQPGGKILVHCAGVSRATLV 159
QY 159 LAYLMHKDMTLDVAIQVAKNRCVLPNRGFLKQLRELDKQLV 202
Db 160 LAYLMYHHLTLVEAIKKVKDHRGIIPNRGFLRLQLALDRRLQ 203

RESULT 3

US-09-164-193-5
Sequence 5, Application US/09164193C
Patent No. 6258582
GENERAL INFORMATION:

APPLICANT: Acton, Susan L.
TITLE OF INVENTION: NOVEL CSAPTP NUCLEIC ACID MOLECULES AND USES THEREFOR (as amended)
FILE REFERENCE: MNI-051
CURRENT APPLICATION NUMBER: US/09/164,193C
CURRENT FILING DATE: 1998-09-30
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 263

TYPE: PRT
ORGANISM: Homo sapiens
US-09-164-193-5

Query Match 31.3%; Score 363; DB 4; Length 263;
Best Local Similarity 51.8%; Pred. No. 3.6e-34;
Matches 73; Conservative 22; Mismatches 44; Indels 2; Gaps 1;

QY 41 LERLEWKGSPQYTHVNEVWPVKLYIGDEATLDRLYRLOKAGFTHTLVNAHGRNVDGTPDY 100
Db 5 VERRLLYTGTACNHADDEVWPGLYLGDQDMANNRRLRLGITHVLNASHSRWR--GTP 62
QY 101 YRDMDIQYHGVEADDLPFDLSVFYPAAAFIDRLSDHSLVHCVMGSRSRATLVLA 160
Db 63 YEGLGIRYLGVEAHDSPAFDMSIHQTAADFTHRLSQPGGKILVHCAGVSRATLVLA 122
QY 161 YLMHKDMTLDVAIQVAKNR 181
Db 123 YLMYHHLTLVEAIKKVKDHR 143

RESULT 4

US-07-988-273-2
Sequence 2, Application US/07988273
Patent No. 5512434
GENERAL INFORMATION:

APPLICANT: AARONSON, Stuart A.
APPLICANT: BOTTARO, Donald P.
APPLICANT: ISHIBASHI, Toshio
APPLICANT: MIKI, Toru
TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
TITLE OF INVENTION: PHOSPHATASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/988,273
FILING DATE: 19921214
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/182 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-988-273-2

Query Match 29.8%; Score 346.5; DB 1; Length 185;
Best Local Similarity 43.8%; Pred. No. 1.8e-32;
Matches 78; Conservative 35; Mismatches 54; Indels 11; Gaps 3;

QY 37 GAFE-----LERLFWKGSPOYT-----HVNVEVWPVKLYIGDEATLDRLYRLOKAGFTHTLVNA 87
Db 3 GSFELSVQDLNDLSDGSGCYSLPQPCNEVTPRIYVGNASVAQDIPKQLKLGITHVLNA 62

Db 63 AEGSRFMHVTNANFYKDSGITYLGKANDTQEFNLSAYFERAADFDQALAKNGRVLV 122
QY 146 HCVMGSRSATLVLAIVLHMKDMLVDALQOQVAKNRCVLPNRCGLKQLRELDKQLVQO 203
Db 123 HCREGYSRPTLVIAVLMRQMDVKLSALSVQRNREIGNPDGFLAQLCOLNDRLAKE 180

RESULT 8
US-07-988-273-3
; Sequence 3, Application US/07988273
; Patent No. 5512434
; GENERAL INFORMATION:
; APPLICANT: AARONSON, Stuart A.
; APPLICANT: BOTTARO, Donald P.
; APPLICANT: ISHIRASHI, Toshio
; APPLICANT: MIKI, Toru
; TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/988,273
; FILING DATE: 19921214
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/182 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 118 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-988-273-3

Query Match 21.3%; Score 247.5; DB 1; Length 118;
Best Local Similarity 47.0%; Pred. No. 3e-21;
Matches 55; Conservative 24; Mismatches 35; Indels 3; Gaps 2;
QY 84 VLNAAGR--WNVDTGPDYRDMIOYHGVEADDLPTDLSVFFYPAA-AFIDRLSDDH 140
Db 1 VLNAAGRSFMHVTNANFYKDSGITYLGKANDTQEFNLSAYFERAADFDQALAKN 60
QY 141 SKILVHCVMGRSRSATLVLAIVLHMKDMLVDALQOQVAKNRCVLPNRCGLKQLRELD 197
Db 61 GRVLVHCREGYSRPTLVIAVLMRQMDVKLSALSVQRNREIGNPDGFLAQLCOLN 117

RESULT 9
PCT-US93-12019-3
; Sequence 3, Application PC/TUS9312019
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12019
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,273
FILING DATE: 14-DEC-1992
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 118 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US93-12019-3

Query Match 21.3%; Score 247.5; DB 5; Length 118;
Best Local Similarity 47.0%; Pred. No. 3e-21;
Matches 55; Conservative 24; Mismatches 35; Indels 3; Gaps 2;
QY 84 VLNAAGR--WNVDTGPDYRDMIOYHGVEADDLPTDLSVFFYPAA-AFIDRLSDDH 140
Db 1 VLNAAGRSFMHVTNANFYKDSGITYLGKANDTQEFNLSAYFERAADFDQALAKN 60
QY 141 SKILVHCVMGRSRSATLVLAIVLHMKDMLVDALQOQVAKNRCVLPNRCGLKQLRELD 197
Db 61 GRVLVHCREGYSRPTLVIAVLMRQMDVKLSALSVQRNREIGNPDGFLAQLCOLN 117

RESULT 10
US-08-990-379-4
; Sequence 4, Application US/08990379
; Patent No. 5998188
; GENERAL INFORMATION:
; APPLICANT: Stork, Philip J
; APPLICANT: Misra-Press, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase CDNAs and
; TITLE OF INVENTION: Their Biologically Active Expression Products
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990,379
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: PCT/US96/10402
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,263
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-08-990-379-4

Query Match 16.6%; Score 192.5; DB 2; Length 393;
Best Local Similarity 34.2%; Pred. No. 4.4e-14;
Matches 50; Conservative 29; Mismatches 58; Indels 9; Gaps 4;
QY 57 EWPKLYIGDEATALDRYRLQKAGFTHVNLAAHGRWNVDTGPDYRDMIOYHGVEADDL 116
Db 197 EILPFLYGSAYHAARDMLDALGTTALLNVS-----SDCPNHF-EGHYQYKCIPIVDN 249
QY 117 PTFDLSVFFYPAAAFIDRLSDDHSHKILVHCVMGRSRSATLVLAIVLHMKDMLVDATQO 176
Db 250 HKADISSWFEAIEYID-AVKDCRGVLVHCQAGISRSATICLAYLMKKRVRLEAEFEF 308
QY 177 VAKNRCVL-PNRCGLKQLRELDKQLV 201
Db 309 VKQRSIIISPNFSPGQLLOFESQVL 334
RESULT 11

RESULT 12 .
US-08-990-379-5
; Sequence 5, Application US/08990379
; Patent No. 5998188

Sequence 5, Application:
Date of: 5000199

Db 174 -EILPYLYGSCNHSDDLQGLQACGITTAVLNVS-----ASCPNHFGL-PHYKSIPIVED 225
Qy 116 LPTDLSVFFYPAAAFIDRALSDHSHKILVHCVMSRSATLVLAYLMIHKDMTLVDAIQ 175
Db 226 NQWEISAWFOEAFISFD-SVKNSGRVLVHCQAGISRSATLCIAYLIQSHRVRLEAFD 284
Qy 176 QVAKNRCVL-PNRGFLKQLRELDKQV 201
Db 285 FVKQRRGVISPNFSGMGLQLLETVL 311

RESULT 14

US-09-164-193-22
; Sequence 22, Application US/09164193C
; Patent No. 6258582
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: NOVEL CSATTP NUCLEIC ACID MOLECULES AND USES THEREFOR (as amended)
; FILE REFERENCE: MNI-051
; CURRENT APPLICATION NUMBER: US/09/164,193C
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 22
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-164-193-22

Query Match 15.9%; Score 184.5; DB 4; Length 314;
Best Local Similarity 28.3%; Pred. No. 2.7e-13;
Matches 56; Conservative 39; Mismatches 84; Indels 19; Gaps 6;

Qy 10 LKNAYSSAKRUSPKMEEGEEDYCTPGAFELER-----LFWKGSQYTHVNEVWPVKLYI 64
Db 128 LRGGFKSFQTCYCPDLCSEAPQAALPPAGAENSNDPRPIYDQGPV-----EILPYLYL 182
Qy 65 GDEATLDRLYRLOKAGFTHVLAAGRWNVDTGPDYRDMDIOYHGVEADDLPFDLSVF 124
Db 183 GSCNHSDDLQGLQACGITTAVLNVS-----ASCPNHFGL-PHYKSIPIVEDNQWVEISAW 235
Qy 125 FYPAAFIDRALSDHSHKILVHCVMSRSATLVLAYLMIHKDMTLVDAIQVAKNRCVL 184
Db 236 FQEAISFD-SVKNSGRVLVHCQAGISRSATLCIAYLIQSHRVRLEAFDFVKQRRGVI 294
Qy 185 -PNRGFLKQLRELDKQV 201
Db 295 SPNFSFGMLQLLETVL 311

RESULT 15

US-09-045-973-5
; Sequence 5, Application US/09045973
; Patent No. 6165767
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah
; TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/09/045,973
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0491 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 198 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTNOT16
CLONE: 3041794
US-09-045-973-5

Query Match 15.3%; Score 177.5; DB 4; Length 198;
Best Local Similarity 28.5%; Pred. No. 8.9e-13;
Matches 53; Conservative 33; Mismatches 61; Indels 39; Gaps 6;

Qy 20 LSPKMEEGEEDYCTPGAFELERLFWKGSQYTHVNEVWPVKLYIGDEATLDRLYRLOKA 79
Db 14 MAPRMISEGD-----IGGIAQITSSSLFLGRGSVASNRHLLQAR 51
Qy 80 GFTHVLNAA-----HGRNVDTGPDYRDMDIOYHGVEADDLPFDLSVFFYPAAAFIDRA 135
Db 52 GITCIVNATIEIPNFN-----POF-----EYKVPPLADMPHAPIGLYFDTVADKI-HS 99
Qy 136 LSDSHSKILVHCVMSRSATLVLAYLMIHKDMTLVDAIQVAKNRCVL-PNRGFLKQLR 194
Db 100 VSRKHGATLVHCAAGVSRSATLCIAYLMKFNHVCLEAYNNVKKARRPVIRENVGVFWRQLI 159
Qy 195 ELDKQL 200
Db 160 DYERQL 165

Search completed: February 8, 2002, 15:42:02
Job time: 121 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:43:46 ; Search time 23.54 Seconds
(without alignments)
692.274 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 220

Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VOQRRSQRQGEEDGREL 220

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size: 10

Total number of hits satisfying chosen parameters: 11

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database: A_Geneseq_1101.*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135	61.4	220	22	Human SGP003 phosph
2	98	44.5	221	21	A human regulator
3	12	5.5	57	22	Human SGP014 phosph
4	12	5.5	80	22	Human polypeptide
5	12	5.5	207	22	Human hydrolyase-li
6	12	5.5	211	21	Human dual specifi
7	12	5.5	211	22	Human polypeptide
8	12	5.5	211	22	Human SGP060 phosph
9	12	5.5	263	21	Human cardiovascular
10	12	5.5	328	22	Human polypeptide
11	12	5.5	549	22	Human SGP014 phosph

ALIGNMENTS

RESULT 1
AAE04837
ID AAE04837 standard; Protein: 220 AA.
XX AC AAE04837;
XX DT 10-SEP-2001 (first entry)
XX DE Human SGP003 phosphatase polypeptide.
XX KW Human: SGP003 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; ocular disease; organ transplant rejection;
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognition disorder; psychotic disorder; cytoskeletal;
KW neurological disorder; viricide; nootropic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; vulnery; tranquiliser; antisthmatic;
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; migraine; chromosome CHR10.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain 1..173
FT FT /label= Catalytic_domain
FT FT 54..199
FT FT /label= Phosphatase_domain
FT FT Misc-difference 85
FT FT /note= "Encoded by CPG"
XX WO200146394-A2.
XX PD 28-JUN-2001.
XX PF 21-DEC-2000; 2000WO-US34736.
XX PR 21-DEC-1999; 99US-0173255.
XX PR 28-DEC-1999; 99US-0175766.
XX PR 25-JAN-2000; 2000US-0178078.
XX PR 31-JAN-2000; 2000US-0179301.
XX PA (SUGE-) SUGEN INC.
XX PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;
XX PI Flanagan P;
XX DR WPI; 2001-418058/44.
XX DR N-PSDB; AAD09495.
XX PT Novel phosphatase polypeptide useful for treating cancers,
XX PT immune-related diseases and disorders, cardiovascular disease, brain or
XX PT neuronal-associated diseases and metabolic disorders
XX PS Claim 7; Fig 2; 186pp; English.
XX CC The present invention relates to phosphatase polypeptides, nucleotide
XX CC sequences encoding them, as well as various products and methods useful
XX CC for the diagnosis and treatment of various phosphatase-related diseases
XX CC and conditions. Substance that modulates the activity of phosphatase
XX CC polypeptide is used to treat immune-related diseases and disorders,
XX CC cardiovascular disease, brain or neuronal-associated diseases and
XX CC metabolic disorders, including cancers of tissues, cancers of
XX CC haematopoietic origin, diseases of central and peripheral nervous
XX CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
XX CC amyotrophic lateral sclerosis, viral infections, infections caused by

CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
 CC sexual dysfunction, mood disorders, attention disorders, cognition
 CC disorders, hypotension, hypertension, psychotic disorders, neurological
 CC amino acid sequence is human SGP003 phosphatase polypeptide. This
 CC sequence is classified as dual specificity phosphatase (DSP) and MAP
 CC kinase phosphatase (MKP). SGP003 gene maps to chromosomal position
 CC CHR10.
 XX
 SQ Sequence 220 AA;

Query Match 61.4%; Score 135; DB 22; Length 220;
 Best Local Similarity 100.0%; Pred. No. 5.3e-128;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 86 NAAHGRNWNVDGPDYRDMDIQYHGVEADLPTFDSLVFFYPAAAFIDRALSDHRSKILV 145
 DB 86 naahgrwnvdtgpyrdmdiqyhgveaddlptfslsvffypaaafidralsdhskilv 145
 QY 146 HCVMGSRSSATLVLAYLMIHKDMTLVDAIQVAKNRCVLPNRGFLKQRLDQLVQQR 205
 DB 146 hcvmgrrsatslvlaylmihkdmtilvdaiqvaknrcvlpnrgflkqireldkqlvqqr 205
 QY 206 RSQRDGGDEEGREL 220
 DB 206 rsqrqgdeedgrel 220

RESULT 2
 AAB18667
 ID AAB18667 standard; Protein; 221 AA.
 AC AAB18667;

22-JAN-2001 (first entry)

A human regulator of intracellular phosphorylation.

KW Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;
 KW neurological disorder; Parkinson's disease; demyelinating disease;
 KW meningitis; developmental disorder; neuromuscular disorder; cancer;
 KW myasthenia gravis; cell proliferative disorder; actinic keratosis;
 KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;
 KW autoimmune disorder; inflammatory disorder; Addison's disease;
 KW acquired immunodeficiency disease; allergy; diabetes mellitus;
 KW rheumatoid arthritis; microbial infection; trauma.

OS Homo sapiens.

Key Location/Qualifiers
 FT Modified-site 2 /note= "potential phosphorylation site"
 FT Modified-site 9 /note= "potential phosphorylation site"
 FT Modified-site 16 /note= "potential phosphorylation site"
 FT Modified-site 21 /note= "potential phosphorylation site"
 FT Modified-site 69 /note= "potential phosphorylation site"
 FT Modified-site 170 /note= "potential phosphorylation site"
 FT Modified-site 208 /note= "potential phosphorylation site"
 FT Active-site 146..158 /note= "tyrosine specific protein phosphatase active site"

W0200055332-A2.

21-SEP-2000.

PF 17-MAR-2000; 2000WO-US07277.
 XX 18-MAR-1999; 99US-0125593.
 PR 20-MAY-1999; 99US-0135049.
 PR 09-JUL-1999; 99US-0143188.
 XX (INCY-) INCYTE PHARM INC.
 PA
 XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;
 PI Lu DAN, Au-Young J;
 PI WPI: 2000-602121/57.
 DR N-PSDB: AAA75684.
 XX Novel intracellular phosphorylation regulator polypeptides and
 FT polynucleotides for diagnosis, prevention and treatment of
 FT neurological, cell proliferative and autoimmune/inflammatory disorders
 FT .
 PT
 PT
 XX Claim 1; Page 84-85; 96pp; English.
 PS
 XX The present sequence represents a human regulator of intracellular
 CC phosphorylation (HRIP). HRIP is useful for screening agonists and
 CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
 CC are useful for treating a disease or condition associated with
 CC decreased or increased expression of functional HRIP. Diseases treated
 CC or diagnosed include neurological disorders such as stroke, Parkinson's
 CC disease, demyelinating diseases, bacterial and viral meningitis and
 CC other developmental disorders of the central nervous system.
 CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
 CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
 CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
 CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
 CC inflammatory disorder such as Addison's disease, acquired
 CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
 CC rheumatoid arthritis, microbial infection and trauma.
 XX
 SQ Sequence 221 AA;

Query Match 44.5%; Score 98; DB 21; Length 221;
 Best Local Similarity 100.0%; Pred. No. 1.1e-90;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTSGEVTSLKNAYSKRLSPKMEEGEEDYCTGAFELERLFWGSPQYTHVNEWVP 60
 DB 1 mtsgevtslknaysakrlspkmeegeedyctgafelerlfwgspqythvnewvp 60
 QY 61 KLYIGDEATLDRLQKAGFTHVLAAGHGRWNVDGTP 98
 DB 61 klyigdeatldrylqkagfthvlnaahgrwnvdtgtp 98

RESULT 3

AAE04849

ID AAE04849 standard; Protein; 57 AA.

AC AAE04849;

10-SEP-2001 (first entry)

Human SGP014 phosphatase polypeptide related exon 3.

KW Human; SGP014 phosphatase polypeptide; phosphatase-related disease;
 KW immune-related disorder; ocular disease; organ transplant rejection;
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
 KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
 KW attention disorder; cognition disorder; psychotic disorder; cyostatic;
 KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;
 KW neuroprotective; antibacterial; vulnery; tranquiliser; antitasthmatic;
 KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;

KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; migraine; chromosome 10q21.3; exon 3.
XX
OS Homo sapiens.
PN WO200146394-A2.
XX
PD 28-JUN-2001.
XX
PF 21-DEC-2000; 2000WO-US34736.
XX
PR 21-DEC-1999; 99US-0173255.
PR 28-DEC-1999; 99US-0175766.
PR 25-JAN-2000; 2000US-0178078.
PR 31-JAN-2000; 2000US-0179301.
XX
PA (SUGEN-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;
PI Flanagan P;
XX
DR WPI: 2001-418058/44.
XX
XX Novel phosphatase polypeptide useful for treating cancers,
PT immune-related diseases and disorders, cardiovascular disease, brain or
PT neuronal-associated diseases and metabolic disorders
XX
PS Example 1; Page 127; 186pp; English.
XX
CC The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorders, including cancers of tissues, cancers of
CC hematopoietic origin, diseases of central and peripheral nervous
CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
CC amyotrophic lateral sclerosis, viral infections, infections caused by
CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognition
CC disorders, hypotension, hypertension, psychotic disorders, neurological
CC disorders, dyskinesias and organ transplant rejection. The present amino
CC acid sequence is human SGP014 phosphatase polypeptide related exon.
CC SGP014 sequence is classified as dual specificity phosphatase (DSP) and
CC MAP kinase phosphatase (MKP). SGP014 gene maps to chromosomal position
CC 10q21.3.
XX
SQ Sequence 57 AA;
Query Match 5.5%; Score 12; DB 22; Length 57;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 152 SRSATFLVLYLM 163
Db 11 srsatflvlylm 22
|||||
RESULT 4
AAM41142
ID AAM41142 standard; Protein; 80 AA.
XX
AC AAM41142;
XX
XX 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6073.
XX
KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB; AAI60298.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries
XX
PS Example 2; SEQ ID NO 6073; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAAM42213) with nontropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, cancer diagnosis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 80 AA;
Query Match 5.5%; Score 12; DB 22; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 152 SRSATFLVLYLM 163
Db 26 srsatflvlylm 37
|||||
RESULT 5
AAB28793
ID AAB28793 standard; Protein; 207 AA.
XX
AC AAB28793;
XX
XX 13-FEB-2001 (first entry)
XX
DE Human hydrolase-like molecule 4 protein.

XX Hydrolase-like molecule; human; cell proliferation disorder;
KW autoimmune; cancer; AIDS; acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN US6132964-A.
XX
PD 17-OCT-2000.
XX
PF 06-FEB-1998; 98US-0013881.
XX
PR 06-FEB-1998; 98US-0013881.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Bandman O, Hillman JL, Guegler KJ, Shah P, Lal P, Corley NC;
XX
DR WPI; 2001-006133/01.
DR N-PSDB; AAC60226.
XX
XX New human hydrolase-like molecules (HHLMs) and polynucleotides encoding
PT the HHLMs, useful for diagnosing, treating or preventing cell
PT proliferation (e.g. bone cancer) or autoimmune disorders(e.g. AIDS or
PT asthma)
XX
PS Claim 1; Column 47-48; 38pp; English.
XX
CC The present invention relates to isolated and purified cDNA encoding a
CC human hydrolase-like molecule (HHLM), designated HHLM-1 to HHLM-8. The
CC HHLM DNAs and polypeptides are useful for diagnosing, treating or
CC preventing cell proliferation disorders and autoimmune disorders. Cell
CC proliferation disorders include cancers and autoimmune disorders include
CC AIDS (acquired immune deficiency syndrome). The present sequence
CC is a HHLM protein of the invention.
XX
SQ Sequence 207 AA;

Query Match 5.5%; Score 12; DB 22; Length 207;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SRSATLVLAYLM 163
Db 153 srsatlvlaylm 164
|||||

RESULT 6
AAV85620
ID AAV85620 standard; Protein; 211 AA.
XX
AC AAV85620;
XX
DT 06-FEB-2001 (first entry)
XX
DE Human dual specificity phosphatase-9 (DSP-9) amino acid sequence.
XX
KW Dual specificity phosphatase-9; DSP-9; human; cytostatic; cell growth;
KW immunosuppressive; antiallergic; antiproliferative; autoimmune disease;
KW cancer; graft-versus-host disease; allergy; metabolic disorder.
XX
OS Homo sapiens.
XX
PN WO200060100-A1.
XX
PD 12-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US09321.
XX
PR 07-APR-1999; 99US-0128203.
XX
PA (CEPT-) CEPTYR INC.

XX Luche RM, Wei B;
XX
DR WPI; 2000-656232/63.
DR N-PSDB; AAC61100.
XX
PT Dual specificity phosphatase-9 which dephosphorylates activated
PT mitogen-activated protein kinase, used to identify agents that inhibit
PT DSP-9 activity and modulate cell proliferation, differentiation, and
PT survival
XX
XX Claim 1; Fig 2A; 66pp; English.
XX
CC This invention relates to an isolated dual specificity phosphatase-9
CC (DSP-9) and its variant. The DSP-9 protein has the ability to
CC dephosphorylate an activated mitogen activated protein (MAP) kinase.
CC Included in the invention are an expression vector comprising a
CC polynucleotide encoding the DSP-9 protein, a host cell transformed by the
CC expression vector, and an antibody that specifically binds to DSP-9.
CC DSP-9 has cytostatic; immunosuppressive; antiallergic; and
CC antiproliferative activity. DSP-9 modulating agents are useful for
CC modulating a proliferative response, differentiation or survival of a
CC cell which displays contacting inhibition of cell growth, anchorage
CC independent growth or an altered intercellular adhesion property, in a
CC patient. DSP-9 agonists and antagonists are also useful for treating a
CC disorder associated with DSP-9 activity such as Duchenne muscular
CC dystrophy, cancer, graft-versus-host disease, autoimmune disease,
CC allergies, metabolic disease, abnormal cell growth, abnormal cell
CC proliferation and cell cycle abnormalities. The present sequence
CC represents the human DSP-9 protein.
XX
SQ Sequence 211 AA;

Query Match 5.5%; Score 12; DB 21; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.00053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SRSATLVLAYLM 163
Db 157 srsatlvlaylm 168
|||||

RESULT 7
AAM39356
ID AAM39356 standard; Protein; 211 AA.
XX
AC AAM39356;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2501.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-052317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.

```

PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR N-PSDB: AAI58512.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries
XX
XX Example 4; SEQ ID NO 2501; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 211 AA;
SQ

```

Query Match 5.5%; Score 12; DB 22; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.00053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 152 SRSATLVLAYLM 163
Db 157 srsatlvlaylm 168
|||||

```

RESULT 8
AAE04839
ID AAE04839 standard; Protein: 211 AA.
XX
XX AAE04839;
AC
XX 10-SEP-2001 (first entry)
DT
XX Human SGP060 phosphatase polypeptide.
DE
XX Human: SGP060 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; ocular disease; organ transplant rejection;
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognition disorder; psychotic disorder; cytostatic;
KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; vulmerary; tranquilliser; antiasthmatic;
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; migraine; chromosome 8p11.1-q11.1.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH

```

Domain 1..173
/label= Catalytic_domain
Domain 61..204
/label= Phosphatase_domain
XX
XX WO200146394-A2.
XX
XX 28-JUN-2001.
XX
XX 21-DEC-2000; 2000WO-US34736.
XX
XX 21-DEC-1999; 99US-0173255.
PR 28-DEC-1999; 99US-0175766.
PR 25-JAN-2000; 2000US-0178078.
PR 31-JAN-2000; 2000US-0179301.
XX
XX (SUGE-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;
PI Flanagan P;
XX
XX WPI: 2001-418058/44.
DR N-PSDB: AAD09497.
XX
XX Novel phosphatase polypeptide useful for treating cancers,
PT immune-related diseases and disorders, cardiovascular disease, brain or
PT neuronal-associated diseases and metabolic disorders
XX
XX Claim 7; Fig 2; 186pp; English.
XX
XX The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase,
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorders, including cancers of tissues, cancers of
CC haematopoietic origin, diseases of central and peripheral nervous
CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
CC amyotrophic lateral sclerosis, viral infections, infections caused by
CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognitive
CC disorders, dyskinesias and organ transplant rejection. The present
CC amino acid sequence is human SGP060 phosphatase polypeptide. This
CC sequence is classified as dual specificity phosphatase (DSP) and MAP
CC kinase phosphatase (MKP). SGP060 gene maps to chromosomal position
CC 8p11.1-q11.1.
XX
XX Sequence 211 AA;
SQ

```

Query Match 5.5%; Score 12; DB 22; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.00053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 152 SRSATLVLAYLM 163
Db 157 srsatlvlaylm 168
|||||

```

RESULT 9
AAAY92175
ID AAY92175 standard; Protein: 263 AA.
XX
XX AAY92175;
AC
XX 01-AUG-2000 (first entry)
DT
XX Human cardiovascular system associated protein tyrosine phosphatase 2.
DE
XX Cardiovascular system associated protein tyrosine phosphatase 2;
KW CSAPTP-2; cytostatic; immunomodulatory; antidiabetic; virucide;

hypotensive; cardiant; tyrosine phosphatase modulator.

Homo sapiens.

Key Location/Qualifiers

Domain 1..131 /label= "N_terminal_domain"

Modified-site 25..30 /note= "unique"

Modified-site 11..13 /note= "N-myristoylation site"

Domain 31..43 /note= "Protein Kinase C phosphorylation site"

Modified-site 47..53 /note= "helix-loop-helix DNA binding domain"

Modified-site 50..53 /note= "tyrosine kinase phosphorylation site"

Modified-site 58..63 /note= "N-glycosylation site"

Domain 76..263 /note= "N-myristoylation site"

Modified-site 89..92 /note= "phosphatase catalytic active domain"

Modified-site 112..117 /note= "Casein kinase II phosphorylation site"

Modified-site 131..134 /note= "N-myristoylation site"

Modified-site 160..165 /note= "Casein kinase II phosphorylation site"

Modified-site 165..168 /note= "N-myristoylation site"

Modified-site 176..179 /note= "Casein kinase II phosphorylation site"

Modified-site 220..222 /note= "amidation site"

Modified-site 223..225 /note= "Protein Kinase C phosphorylation site"

Modified-site 246..251 /note= "Protein Kinase C phosphorylation site"

Modified-site 257..260 /note= "N-myristoylation site"

Modified-site 257..260 /note= "Casein kinase II phosphorylation site"

WO200018890-A2.

06-APR-2000.

30-SEP-1999; 99WO-US22924.

30-SEP-1998; 98US-0163833.

30-SEP-1998; 98US-0164193.

(MILL-) MILLENNIUM PHARM INC.

Acton S;

WPI; 2000-293136/25.

N-PSDB; AAA09027, AAA09028.

New cardiovascular system associated protein tyrosine phosphatases are

used in the diagnosis and treatment of e.g. immune disorders,

anti-proliferative disorders, metabolic disorders and cardiovascular

disorders

Claim 8; Page 146-147; 156pp; English.

The CSATPP nucleic acid and protein molecules are used to modulate

regulation of cellular processes. CSATPP nucleic acid and protein

molecules and modulators of CSATPP activity and expression can be

used to treat a subject with a disorder characterized by aberrant CSATPP

expression or activity. These disorders can include an immune disorder,

an anti-proliferative disorder, a proliferative disorder e.g. renal and

lung carcinomas, a metabolic disorder e.g. diabetes, viral pathogenesis,

CC a neural disorder, a cardiovascular disorder e.g. hypertension or
CC coronary heart disease or a disorder rising from improper phosphorylation
CC of a phosphorylated protein. Anti-CSATPP antibodies are used to isolate
CC CSATPP by standard techniques, to facilitate the purification of natural
CC and recombinantly produced CSATPP from cells, to detect CSATPP protein in
CC cell supernatant for evaluating the level of CSATPP expression and to
CC monitor protein levels in tissue as part of a clinical testing procedure
CC to determine efficacy of a treatment procedure. CSATPP nucleic acids, and
CC antibodies and protein molecules can be used in screening assays and in
CC predictive medicine e.g. prognostic assays, monitoring clinical trials
CC and pharmacogenetics.

xx Sequence 263 AA;

Query Match 5.5%; Score 12; DB 21; Length 263;

Best Local Similarity 100.0%; Pred. No. 0.00065;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 152 SRSATLVLAYLM 163

Db 114 srsatlvlaylm 125

RESULT 10

AA42355

ID AA42355 standard; Protein; 328 AA.

XX AA42355;

AC AA42355;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 88.

XX

Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;
immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
anticarionomic; antisickling; antianaemic; antithrombotic; cancer;
antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
antiparasitic; cardiant; immune disorder; cardiovascular disorder;
neurological disease; infection; nephrotropic; gene therapy; vaccine.

OS Homo sapiens.

XX WO200155449-A1.

PN 02-AUG-2001.

PD 17-JAN-2001; 2001WO-US01346.

PF 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUL-2000; 2000US-0216880.

PR 14-JUL-2000; 2000US-0218290.

PR 14-AUG-2000; 2000US-0225447.

PR 01-SEP-2000; 2000US-0229343.

PR 06-SEP-2000; 2000US-0230437.

PR 08-SEP-2000; 2000US-0231243.

PR 25-SEP-2000; 2000US-0234997.

PR 29-SEP-2000; 2000US-0236367.

PR 13-OCT-2000; 2000US-0239937.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 17-NOV-2000; 2000US-0246528.

PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249265.

PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SG, Ruben SM;
 XX WPI; 2001-476225/51.
 XX N-PSDB; AAI62760.

PT Novel plasma membrane associated proteins useful for diagnosing,
 PT treating, preventing and/or prognosing disorders related to the
 PT proteins, including cancer, immune response and neuronal disorders

PS Claim 11: SEQ ID NO 88; 532pp + Sequence Listing; English.

XX The invention relates to novel genes (AAI62752-AAI62961) and proteins
 CC (AAW42347-AAW42415) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 328 AA;

Query Match 5.5%; Score 12; DB 22; Length 328;
 Best Local Similarity 100.0%; Pred. No. 0.00079;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 SRSATLVLYLML 163
 Db 274 srsatlvlylm 285
 |||||

RESULT 11
 AAE04838

ID AAE04838 standard; Protein: 549 AA.

AC AAE04838;

DT 10-SEP-2001 (first entry)

DE Human SGP014 phosphatase polypeptide.

XX Human; SGP014 phosphatase polypeptide; phosphatase-related disease;
 KW immune-related disorder; ocular disease; organ transplant rejection;
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
 KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
 KW attention disorder; cognition disorder; psychotic disorder; cytostatic;
 KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;
 KW neuroprotective; antibacterial; vulnery; tranquiliser; antasthmatic;
 KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
 KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;

KW MKP; migraine; chromosome 10q21.3.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..173
 FT FT /label= Catalytic_domain
 FT Domain 37..181
 FT FT /label= Phosphatase_domain
 FT Domain 368..520
 FT FT /label= Phosphatase_domain

XX WO200146394-A2.

XX 28-JUN-2001.

XX 21-DEC-2000; 2000WO-US34736.

XX 21-DEC-1999; 99US-0173255.

PR 28-DEC-1999; 99US-0175766.

PR 25-JAN-2000; 2000US-0178078.

PR 31-JAN-2000; 2000US-0179301.

XX (SUGE-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;

PI Flanagan P;

XX WPI; 2001-418058/44.

DR N-PSDB; AAD09496.

XX Novel phosphatase polypeptide useful for treating cancers,
 PT immune-related diseases and disorders, cardiovascular disease, brain or
 PT neuronal-associated diseases and metabolic disorders

XX Claim 7; Fig 2; 186pp; English.

XX The present invention relates to phosphatase polypeptides, nucleotide
 CC sequences encoding them, as well as various products and methods useful
 CC for the diagnosis and treatment of various phosphatase-related diseases
 CC and conditions. Substance that modulates the activity of phosphatase
 CC polypeptide is used to treat immune-related diseases and disorders,
 CC cardiovascular disease, brain or neuronal-associated diseases and
 CC metabolic disorders, including cancers of tissues, cancers of
 CC haematopoietic origin, diseases of central and peripheral nervous
 CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
 CC amyotrophic lateral sclerosis, viral infections, infections caused by
 CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
 CC sexual dysfunction, mood disorders, attention disorders, cognition
 CC disorders, dyskinesias, hypertension, psychotic disorders, neurological
 CC disorders, and organ transplant rejection. The present
 CC amino acid sequence is human SGP014 phosphatase polypeptide. This
 CC sequence is classified as dual specificity phosphatase (DSP) and MAP
 CC kinase phosphatase (MKP). SGP014 gene maps to chromosomal position
 CC 10q21.3.

XX Sequence 549 AA;

Query Match 5.5%; Score 12; DB 22; Length 549;
 Best Local Similarity 100.0%; Pred. No. 0.0013;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 SRSATLVLYLML 163

Db 134 srsatlvlylm 145
 |||||

Search completed: February 8, 2002, 15:45:46
 Job time: 120 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:45:01 ; Search time 15.43 Seconds
(without alignments)
1086.092 Million cell updates/sec

Title: US-09-847-519A-2
Perfect score: 220
Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VOORRRSQRDGEEDGREL 220

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 10

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Query Score	Match Length	DB ID	Description
---------------	----------------	-----------------	-------	-------------

No matches found

Search completed: February 8, 2002, 15:48:10
Job time: 189 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:46:07 ; Search time 11.72 seconds
(without alignments)
688.248 Million cell updates/sec

Title: US-09-847-519A-2
Perfect score: 220
Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VQRRRSQRQDGEEDGREL 220

Scoring table: OLIGO
Gapop 60.0' , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 10

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
-----	-----	-----	-----	-----	-----

No matches found

Search completed: February 8, 2002, 15:49:20
Job time: 193 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:45:47 ; Search time 23.66 seconds
(without alignments)
1360.098 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 220

Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VQRRRSQRQDGEEDGREL 220

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 10

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	5.5	86	4 Q9BTW0	Q9btw0 homo sapien
2	12	5.5	198	11 Q9D700	Q9d700 mus musculu
3	12	5.5	211	4 Q9BV47	Q9bv47 homo sapien

ALIGNMENTS

RESULT 1

Q9BTW0

ID: Q9BTW0 PRELIMINARY; PRT: 86 AA.

AC Q9BTW0:

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE UNKNOWN (PROTEIN FOR MGC:2627).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=NEUROBLASTOMA;

RA Strausberg R.;

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC003115; AA031115.1;

SQ SEQUENCE 86 AA; 9652 MW; 4DD29AFD5989528B CRC64;

Query Match 5.5%; Score 12; DB 4; Length 86;

Best Local Similarity 100.0%; Pred. No. 0.00017;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 SRSATLVLYLML 163

|||||

Db 32 SRSATLVLYLML 43

RESULT 2

Q9D700 PRELIMINARY; PRT: 198 AA.

AC Q9D700;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DE 2310043K02RIK PROTEIN.

GN 2310043K02RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=TONGUE;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Kasaiwa H.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

Nature 409:685-690(2001).

DR EMBL; AK009781; BAB26501.1;

DR MGD; MGI:1914209; 2310043K02RIK.

DR InterPro; IPR000340; DS_phosphatase.

DR InterPro; IPR000387; TYR_phosphatase.

DR Pfam; PF00782; DSPC; 1.

DR SMART; SM00195; DSPC; 1.

DR SMART; SM00012; PTPC_DSPC; 1.

DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.

DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.

DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.

KW Hydrolase.

SQ SEQUENCE 198 AA; 22419 MW; 11F0CAF4B5620F0E CRC64;

Query Match

Best Local Similarity 100.0%; Score 12; DB 11; Length 198;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SRSATLVLAYLM 163
 |||||
 Db 144 SRSATLVLAYLM 155

RESULT 3

Q9BV47 PRELIMINARY; PRT; 211 AA.
 AC Q9BV47;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE UNKNOWN (PROTEIN FOR MGC:1136).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG CARCINOMA;
 RA Strausberg R.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC001613; AAH01613.1; -;
 SQ SEQUENCE 211 AA; 23945 MW; 60E944304905086D CRC64;

Query Match 5.5%; Score 12; DB 4; Length 211;
 Best Local Similarity 100.0%; Pred. No. 0.00037;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SRSATLVLAYLM 163
 |||||
 Db 157 SRSATLVLAYLM 168

Search completed: February 8, 2002, 15:49:02
 Job time: 195 sec

Result No.	Query	Score	Match	Length	DB	ID	Description
1	199.4	17.1	594	4	US-09-163-833-3	Sequence 3, Appli	
2	199.4	17.1	912	4	US-09-163-833-1	Sequence 1, Appli	
3	175	15.0	1691	3	US-09-013-881-12	Sequence 12, Appl	
4	161.6	13.9	789	4	US-09-164-193-6	Sequence 6, Appli	
5	161.6	13.9	1016	4	US-09-164-193-4	Sequence 4, Appli	
6	133	11.4	861	1	US-07-988-273-1	Sequence 1, Appli	
7	133	11.4	861	5	PCT-US93-12019-1	Sequence 1, Appli	
8	70.6	6.1	1238	2	US-08-530-290-11	Sequence 11, Appl	
9	60.2	5.2	1987	2	US-08-990-379-1	Sequence 1, Appli	
10	60.2	5.2	1993	2	US-08-990-379-2	Sequence 2, Appli	
11	51.6	4.4	4451	3	US-08-717-294-42	Sequence 42, Appl	
12	49.8	4.3	2097	3	US-08-941-445A-10	Sequence 10, Appl	
13	48.6	4.2	1176	2	US-08-387-942C-17	Sequence 17, Appl	
14	48.6	4.2	12588	2	US-08-387-942C-1	Sequence 1, Appli	
15	48.2	4.1	11219	1	US-07-642-734C-1	Sequence 1, Appli	
16	48.2	4.1	11219	3	US-08-439-009A-1	Sequence 1, Appli	
17	47.6	4.1	1729	4	US-09-045-973-6	Sequence 6, Appli	
18	47.2	4.1	7218	1	US-08-232-463-14	Sequence 14, Appl	
19	47	4.0	1704	1	US-08-528-199-2	Sequence 2, Appli	
20	47	4.0	1704	1	US-08-528-199-5	Sequence 5, Appli	
21	46.2	4.0	5515	4	US-09-398-193-98	Sequence 98, Appl	
22	46.2	4.0	30001	1	US-08-125-468-1	Sequence 1, Appli	
23	46.2	4.0	30001	2	US-08-474-933-1	Sequence 1, Appli	
24	45.8	3.9	1557	5	PCT-US91-01327-12	Sequence 12, Appl	
25	45.8	3.9	1557	6	5248670-2	Sequence 12, Appl	
26	45.6	3.9	2943	1	US-08-042-747A-7	Patent No. 5248670	
27	44.8	3.8	2825	4	US-09-196-390-5	Sequence 7, Appli	

Db	826	CAGCCAGGAGGAAGATCCTGCTGTCATTGCTGTGGCGGTGAGCCGATCCGCCACCCCTG	885
Qy	591	gtcctggcctacctgatgatccacaagacatgacccctggctggcgcacatccagcaagtg	650
Db	886	GTACTGGCCTACTCATGCTGTACCAACCACCTTACCTTCGTGGAGGCCATCAAGAAATC	945
Qy	651	gccaaagaacgctgcgctctcccgaccggggcctttttaagcagctcccgagctggac	710
Db	946	AAAGACCACGAGGATCATCCCCAACCGGGGCTTCCTGAGGCAGCTCCTGGCCCTGGAC	1005
Qy	711	aagcagctggtgcag	725
Db	1006	CGCAGGCTGCGCAG	1020

```

RESULT      4
US-09-164-193-6
; Sequence 6, Application US/09164193C
; Patent No. 6258582
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: NOVEL CSATTP NUCLEIC ACID MOLECULES AND USES THEREFOR (as amended)
; FILE REFERENCE: MNI-051
; CURRENT APPLICATION NUMBER: US/09/164,193C
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 6
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(789)
US-09-164-193-6

```

Query Match	13.9%;	Score 161.6;	DB 4;	Length 789;
Best Local Similarity	61.9%;	Pred. No. 4.5e-28;		
Matches 276;	Conservative 0;	Mismatches 164;	Indels 6;	Gaps
QY 245	gcggctctcttgaaaggcagtcgccaggtacaccacgctcaacagaggtctgcccacagct	304		
Db 18	gcggctctcttacacaggaacagacctgtaacctgcgcagaggtctgcccagacct	77		
QY 305	ctacattggcgatgagcgacgcgcgtgaccgcgtataggctgcagaaagcgagggttcac	364		
Db 78	ctatctcgagacacagacatggctaacacgcgcggagcttcgcgcgtgggcacac	137		
QY 365	gcagctgctgaacgcgcgcacgcgcctggaaagtggacactggcccgactactacag	424		
Db 138	gcagctctctcaatgctctcacacagcgggtggcgag-----gcacgcccagagcctatga	191		
QY 425	cgacatggacatcagtatcaacggcctggagggccgacacctgccacactctgcacctcag	484		
Db 192	ggggctgggcatcgcgtactcgtggtgtgaggccacgactgcgcagcctttgacatgag	251		
QY 485	tgctctctctacccggcggcagccttcctatgcagagcgctaaagcgacgacacacagtaa	544		
Db 252	catccacttcacagcgtgcgcacttcctacaccggcgctgagccagccagagggaag	311		
QY 545	gatccttggttacctgcgtcatgggcgcgacgcggtcagccacctggtcctggcctacct	604		
Db 312	gatccttggtgcatgtgctgtgggcgtgagcgatccgcacacctggttactggcctacct	371		
QY 605	gatgtccacaagaagcatgaccttggtgacgcctatccagcaagtggccaaagaacctg	664		
Db 372	catgtgtaccaccacttacctctcgtggaggccaacgaagaagtcaagaccaccgag	431		
QY 665	cgctctcccgaaacggggctttttga	690		
Db 432	agagggccgagcccagccactatca	457		

```

: RESULT 6
: US-07-988-273-1
: ; Sequence 1, Application US/07988273
: ; Patent No. 5512434
: ; GENERAL INFORMATION:
: ; APPLICANT: AARONSON, Stuart A.
: ; APPLICANT: BOTTARO, Donald P.
: ; APPLICANT: ISHIBASHI, Toshio
: ; APPLICANT: MIKI, Toru
: ; TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
: ; TITLE OF INVENTION: PHOSPHATASE
: ; NUMBER OF SEQUENCES: 7
: ; CORRESPONDENCE ADDRESS:

```


APPLICANT: Hughes, David Anthony
TITLE OF INVENTION: Methods for Screening of Substances for
TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,290
FILING DATE: 14-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB94/00694
FILING DATE: 31-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9402573.1
FILING DATE: 10-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9307250.2
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 084611-00000005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1238 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-530-290-11

Query Match 6.1%; Score 70.6; DB 2; Length 1238;
Best Local Similarity 52.9%; Pred. No. 2.3e-07;
Matches 176; Conservative 0; Mismatches 154; Indels 3; Gaps 1;
QY 393 tgaacgtgacactggccgactactaccgacatgacatccagaccgagcgtg 452
Db 729 TTGAATGTCCTCGGACTGCCAACCCTTTGAAGGACACTATCAGTCAAGTCGATC 788
QY 453 gagcgacgacactggccacacttgcacactcagtgctcttcttaccggcgagccttc 512
Db 789 CCACTGGAAGATACCAACCAAGCGGACATCAGCTCCTGTTTCATGGAAGCCATAGATAC 848
QY 513 atcgacagagcgttaagcagcagaccacagtaagatcctggttactcgtcgtatggcgcg 572
Db 849 ATCGA---TCCCGGTGAAGGACTGCCGTGGCGCGCTGCTGTGTCACATGCCAGCGGGCATC 905
QY 573 agcgcggtcagccaccctggcttggcctactgatgatccacagacatgacacctggtg 632
Db 906 TCGCGGTGCGCCACCATCTGCTGGCCTACCTGATGATGATGATGATGATGATGATGATGATG 965
QY 633 gagccatccagcaagtgcccaaacctgctgctcctcccgaaaccgggggctttttgaag 692
Db 966 GAGGCCCTTCGAGTTCGTTAAGCAGCGCGGACGATCATCTCGCCCACTTCAGCTTCATG 1025
QY 693 cagctcgggagctggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 725
Db 1026 GGCAGCTGCTGCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG

RESULT 9

US-08-990-379-1
Sequence 1, Application US/08990379
Patent No. 5998188
GENERAL INFORMATION:
APPLICANT: Stork, Philip J
APPLICANT: Misra-Press, Anita
TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
TITLE OF INVENTION: Their Biologically Active Expression Products
FILE REFERENCE: 4104-000322USA
CURRENT APPLICATION NUMBER: US/08/990,379
CURRENT FILING DATE: 1997-12-15
EARLIER APPLICATION NUMBER: PCT/US96/10402
EARLIER FILING DATE: 1996-06-14
EARLIER APPLICATION NUMBER: 60/000,263
EARLIER FILING DATE: 1995-06-16
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1987
TYPE: DNA
ORGANISM: Rattus norvegicus
US-08-990-379-1

Query Match 5.2%; Score 60.2; DB 2; Length 1987;
Best Local Similarity 51.1%; Pred. No. 6.2e-05;
Matches 168; Conservative 0; Mismatches 158; Indels 3; Gaps 1;
QY 395 gaacgtggacactggccgactactaccgacatgacatccagaccgagcgtgga 454
Db 1052 gaatgtctctcagactgcccccaatcactttgaggagacattaccagtcacagtcacccc 1111
QY 455 ggcgcgacgacctgccacacttgcacactcagtgcttcttcttaccggcgagccttcac 514
Db 1112 ggtagaagataaccacaaagctgacatcagctcctggttcattggaagccatcgaatacat 1171
QY 515 cgacagagcgttaagcagcagcaccacagtaagatcctggttactcgtcagtcagtcagcag 574
Db 1172 agac---gcagtgaaggagactggcgagtgctggttactcagtcagtcagtcagtcagtc 1228
QY 575 ccggtcagccaccctgctcctggcctacctgatgatccacagacatgacacctggtgga 634
Db 1229 tagatccgcccaccatctgcttggcctacctgatgatgaagaacgggtgaggtctggaaga 1288
QY 635 cgccatccagcaagtggccaaagacccgctgctcctcccgaaaccggggctttttgaagca 694
Db 1289 ggcttgcagttcgtcaagcagcgcgtagcatcatctcgcaccaacttcagcttcatggg 1348
QY 695 gctcgggagctggacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 723
Db 1349 ccagttgctgcagttcgcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag

RESULT 10

US-08-990-379-2
Sequence 2, Application US/08990379
Patent No. 5998188
GENERAL INFORMATION:
APPLICANT: Stork, Philip J
APPLICANT: Misra-Press, Anita
TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
TITLE OF INVENTION: Their Biologically Active Expression Products
FILE REFERENCE: 4104-000322USA
CURRENT APPLICATION NUMBER: US/08/990,379
CURRENT FILING DATE: 1997-12-15
EARLIER APPLICATION NUMBER: PCT/US96/10402
EARLIER FILING DATE: 1996-06-14
EARLIER APPLICATION NUMBER: 60/000,263
EARLIER FILING DATE: 1995-06-16
NUMBER OF SEQ ID NOS: 19


```

1  RESULT 14
2  US-08-387-942C-1
3  US-08-387-1, Application US/08387942C
4  Sequence 1, Patent No. 5939289
5  Patent No. 5939289
6  GENERAL INFORMATION:
7  APPLICANT: ERTESVAG, HELGA
8  APPLICANT: VALLA, SVEIN
9  APPLICANT: SKJAK-BRAEK, GUDMUND
10 APPLICANT: LARSEN, BJORN
11 TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES
12 OF INVENTION: ENCODING MANNURONAN C-5-EPIMERASE
13 NUMBER OF SEQUENCES: 52
14 CORRESPONDENCE ADDRESS:
15 ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
16 STREET: P.O. BOX 747
17 CITY: FALLS CHURCH
18 STATE: VA
19 COUNTRY: USA
20 ZIP: 22042
21 COMPUTER READABLE FORM:
22 MEDIUM TYPE: Floppy disk

```

TITLE OF INVENTION: Erythromycin Analogs
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman
STREET: Abbott Laboratories D377/AP6D-2 One Abbott
STREET: Park Rd
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/642,734C
FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dancckers, Andreas M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11219 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharopolyspora erythraea
STRAIN: NRRL 2338
FEATURE:

```

FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 744..6659
  OTHER INFORMATION: /function= "APPROXIMATE SPAN OF
  OTHER INFORMATION: MODULE 1"
  OTHER INFORMATION: /label= FUNCTION
FEATURE:
  NAME/KEY: CDS
  LOCATION: 744..11219
  OTHER INFORMATION: /function= "gene= "eryA""
  OTHER INFORMATION: /product= "ORF1 encoding modules 1 & 2 for
  OTHER INFORMATION: 6-deoxyerythronolide B""
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 744..1868
  OTHER INFORMATION: /function= "approximate span of
  OTHER INFORMATION: acyltransferase domain 1 of module 1"
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 1998..2198
  OTHER INFORMATION: /function= "approximate span of
  OTHER INFORMATION: acyl carrier domain 1 of module 1"
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 2250..3626
  OTHER INFORMATION: /function= "approximate span of
  OTHER INFORMATION: beta-ketoacylACP synthase domain/module1"
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 3831..4811
  OTHER INFORMATION: /function= "approximate span of
  OTHER INFORMATION: acyltransferase 2 domain of module 1"
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 5574..6125

```

```

FEATURE:
  NAME/KEY:  misc_feature
  LOCATION:  3831..4811
  OTHER INFORMATION:  /function= "approximate span of
  OTHER INFORMATION:  beta-ketoacylACP synthase domain/module1"
FEATURE:
  NAME/KEY:  misc_feature
  LOCATION:  5574..6125
  OTHER INFORMATION:  /function= "approximate span of
  OTHER INFORMATION:  acyltransferase 2 domain of module 1"

```

```
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6369..6826
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain 2 of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6678..11219
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6678..8066
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl ACPsynthase of module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8262..9305
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9906..10454
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10707..10964
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 2"
US-07-642-734C-1
```

```
Query Match 4.1%; Score 48.2; DB 1; Length 11219;
Best Local Similarity 45.9%; Pred. No. 0.057;
Matches 164; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

QY 362 cagcagcgtgtaagcagcgccgacgctggaacgtggaacacgtggccgactacta 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9950 CGCGAGGTGGCCCGGTGGCTCCCGGTGTCGCGCGCGGACACCTCGCGTGTGTCAGCAG 10009

QY 422 ccgacatggacatccagtacacagcggtggagggcagacgtgccaccttcgacct 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10010 GCGAGGCCGACACCGAGGGCGTCCGACCTGACCGCGAGCTGACCGCGCTCGGCGC 10069

QY 482 cagtgcttcttaccggcgagccttcctcctcagacagcgtcgaagcagaccacag 541
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10070 GCGGGTGTGCGTGCACGCGTGCAGCGTACGACGCGGACCGGACCGTGTGAGGAAC 10129

QY 542 taagatcctggttcactgctgctggtggcgagcagcgtcagccacctgctgacct 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10130 CGGCCCTGATCGACAGGGCGACGTCGTCGCGGTGTGTCACGCGCGGACTGCGGCA 10189

QY 602 cctgatgatccaaagacatgacctggtggagcgccatccagcagcagtgccaaagaccg 661
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10190 GCAGGTGCGGATCACGACATGACGAGGCGCGCTTCGACGAGGTGCTCGCGCCAGGC 10249

QY 662 ctgctctctccgaacccggggtttttgaagcagctccgggagctggacaagcagct 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10250 CGGGGGCGCGGTGCACCTGGACGAGTGTGTCGGAGCGCGAGCTGTTCTGCTGTT 10306
```

Search completed: February 8, 2002, 18:43:16
Job time: 4241 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2002, 18:14:05 ; Search time 1166.24 Seconds
(without alignments)
10734.353 Million cell updates/sec

Title: US-09-847-519A-1
Perfect score: 1165
Sequence: 1 ggcagtggtggtggtggg.....aaaaaaaaaaaaaaaaaaaaa 1165

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	444.4	38.1	572	11	BF074326
2	425.8	36.5	634	10	AW918423
3	401	34.4	605	10	AW918455
4	368	31.6	478	11	BF652341
5	242.4	20.8	505	11	BE808287
6	200	17.2	486	11	BF430044
7	194.4	16.7	703	11	BI199835
8	190	16.3	1748	12	AK009781
9	187.6	16.1	558	11	BF077053
10	176.6	15.2	990	11	BF207232
11	166.2	14.3	1104	12	AK007513
12	165.8	14.2	471	11	BF774179

13	164	14.1	415	10	AI372800
14	163.6	14.0	1065	12	AK006247
15	161.8	13.9	456	10	BE751020
16	160.4	13.8	612	11	BI197674
17	159.8	13.7	469	10	AW425509
18	158.2	13.6	961	11	BF314818
19	153.6	13.2	693	11	BG083406
20	148.8	12.8	369	11	BF430184
21	148.4	12.7	607	10	AL585353
22	144.2	12.4	411	11	BF826456
23	143.6	12.3	641	11	BI393955
24	135.6	11.6	642	11	BI393954
25	134.4	11.5	1149	11	BI083188
26	133	11.4	550	11	BE892660
27	133	11.4	644	11	BG17129
28	133	11.4	723	10	BE387475
29	133	11.4	850	11	BE750047
30	133	11.4	949	10	BE383984
31	133	11.4	1001	10	AL555009
32	131.6	11.3	561	10	AL602266
33	129.6	11.1	461	11	BF776211
34	129.4	11.1	566	10	AI526580
35	128.8	11.1	946	11	BG246446
36	128.8	11.1	1169	12	AK008734
37	127.2	10.9	718	10	BE390181
38	125.8	10.8	545	10	BE757502
39	121.4	10.4	957	11	BF527844
40	118	10.1	713	11	BF212271
41	116.4	10.0	781	10	AU168945
42	114.8	9.9	609	10	AW951243
43	114.8	9.9	684	10	BE613888
44	113.2	9.7	430	10	BE771213
45	111.4	9.6	927	10	BE728819

ALIGNMENTS

RESULT 1
BF074326 LOCUS BF074326 572 bp mRNA EST 25-APR-2001
DEFINITION 221711 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BF074326
VERSION BF074326.1 GI:10867837
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 572)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett ,G.L., Heaton,M.P., Laegreid,W., Rohrer,G.A., Chalko-McKown,C.G., Portea,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGACG
Plate: 80 row: I column: 23


```

Bovidae; Bovinae; Bos.
1 (bases 1 to 486)
Warren,W.C., Tao,N., Allison,T., Wagner,S., Mathialagan,N., Kata,S.,
Johnson,J., Smith,T.P. and Womack,J.
A survey of genes transcribed in bovine skeletal muscle
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel.: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR primers
FORWARD: GGAAACAGCTATGACCATG
BACKWARD: GTTTCCTCAGTCACGAC
Seq primer: AATTAACCCCTCACTAAAGGG.
Location/Qualifiers
1..486
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC BSM"
/tissue_type="Skeletal muscle"
/lab_host="XLOLR"
/note="Vector: Uni-ZAP.XR; Site_1: EcoRI; Site_2: XhoI;
Library obtained from Stratagene, catalog #937721. Library
made from skeletal muscle of a two year old Holstein cow."
96 a 147 c 136 g 107 t

```

```

DEFINITION      602761093F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4896479 5',
                mRNA sequence.
ACCESSION      BI199835
VERSION        BI199835.1 GI:14654856
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 703)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-r@mail.nih.gov
                Tissue Procurement: ATCC
                CDNA Library Preparation: Ling Hong/Rubin Laboratory
                CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
                Plate: LLC1M783 row: 1 column: 24
                High quality sequence stop: 701.

FEATURES             Location/Qualifiers
     source           1..703
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4896479"
                     /clone_lib="NIH_MGC_19"
                     /tissue_type="neuroblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Library constructed by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-CDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: This is a NIH_MGC Library."
BASE COUNT       139 a 223 c 223 g 118 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
89
```



```
OY 374 gaacgagccacgagcgcgtgaacgtgacactgagccgactactaccgcgacatgga 433
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 729 CAACGCTCACACACAGTGGGAGGCACC-----CCGAGGCTATGAGGACTGG 782

OY 434 catcagatcacacgagcgtgagggccgacgacctgcccacttcgacctcagtgcttctt 493
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 783 CATCGCTACCTGGTGTGGAGGGCCAGCACTCTCCAGCTTTGATATGACATCCACTT 842

OY 494 ctaccgagcgcagccttcacgcagacagcgttaagcagcagcaccacataagatcctagt 553
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 843 TCAGAGGCTGGGACTTCATCCATCGGGCGCTGAGCCAGCAGGAGGAAGATCCTGCT 902

OY 554 tcaatgcctcatggggccgagcgcgtgagccacacctgctcctgacctgatgatcca 613
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 903 GCAGTGTGGGAGTGAGCAGATCCGCTACCTGCTGCTGGCTACCTCATCTGTA 962

OY 614 caagagacatgacctggtgagcgcacatccagcaagtgccaaagaccctgcgtcctccc 673
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 963 TCACCACCTTACCTCTGTGGAGGGCCATCAAGAAAGCTCAAGGACCACCGAGGCATCACCCC 1022

OY 674 gaaccgggagcctttgaagcagcgtccggagctggacagcagcagcgtggtgcag 725
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1023 CAACCGGGGCTCTTCAGCGAGCTCTAGCCCTGAGCCAGTGGCTGCGGCAG 1074

RESULT 9
BF077053
LOCUS BF077053 558 bp mRNA EST 25-APR-2001
DEFINITION 226847 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BF077053
VERSION BF077053.1 GI:10870883
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 558)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G.,
Perte,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keele,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the "minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 89 row: A column: 6
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1..558
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
96 a 180 c 171 g 111 t

BASE COUNT
```

```
ORIGIN
Query Match 16.1%; Score 187.6; DB 11; Length 558;
Best Local Similarity 61.7%; Pred. No. 3.8e-24;
Matches 317; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

OY 201 gggagggagggagggactactgcacccctggagcctttgagctggagcgtctcttcggaag 260
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 48 GGAGATGCGGAGCCACGCCCTTGCCCCAGCGTCTGGGAACCTGGAGGAGCTCTCTGAGGCA 107

OY 261 ggcagtcctccagttacacccagtcgaacgtgctggcccaagctctacatggcgatgag 320
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 108 GGAAGGTTTCTTCCAGCCACGCTGGATGAAGTTTGGCCCAACCTTTACATAGGAGATGCG 167

OY 321 gcgaggggagctggagcgcctataggctgcagaagcgggggttcacgcagctgctgaacgcg 380
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 168 GCCACAGCAATAAACCGCTTTGAGCTATGGAAGCTGGGCATTTACCCAGCTGCTGAATGCC 227

OY 381 gccacggcgcctggagcgtggaacactggcccgactactaccgcgacatggacatccag 440
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 228 GCCACAGGGGGCTCTACTGTTCAGGGCAGCCCTGACTTCTATGGCAGCAGTGTGAGC--- 284

OY 441 taccacggcgtggagggcgcgacgtgccaccccttcgacctcagtgctcttcttaccgcg 500
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 285 TACCTGGGGTGGCCAGCCACGACCTTCTTGAGTTTCGACATCAGTGTCTACTTCTCTCT 344

OY 501 gcggcagccttcacgcagcagcagcgtgtaagcgcagcaccacagtaagatccttggtcactgc 560
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 345 GCAGTGACTTTCATCCACGCTGCCCTTAGCAGCGCTGGGGCCAGAGGTCTGCTGTCACATGC 404

OY 561 gtcagggcgcgagcgcgtcagcaccctgctcctggcctaccctgctgacacacagcag 620
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 405 GTGTGCGGGGTGAGCGGCTCCGCCACACTGGTCTCTGGGCTTACTCATGCTGCGCCAGCAG 464

OY 621 atgacccgtgtgagcgcattccagcaagtggcgaagcagcgtgctcctccgcgaacgg 680
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 465 CTCCTCCCTGCCAGCGGTGATCACCCTGAGGAGCGCCGATGGTCTTCCCAACAGA 524

OY 681 ggcctttgagcagcgtccggagcgtgacaaagc 714
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 525 GGTTCCTTCCACAGCTCTGCGGCTGGACCAGC 558

RESULT 10
BF07232
LOCUS 601870679F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4100947 5',
DEFINITION mRNA sequence.
ACCESSION BF07232
VERSION BF07232.1 GI:11100818
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 990)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L10M971 row: n column: 20
High quality sequence stop: 582.
Location/Qualifiers
1..990
/organism="Homo sapiens"

FEATURES
source
```


Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S. Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dinke, D., Feng, D.-F., Ferlie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL
Nature 377 (8547 Suppl), 3-174 (1995)

MEDLINE
96026280

COMMENT
Other ESTs: EST175329 EST175330 EST175331 EST175332

EST175333 EST175334 EST175335 EST175336 EST175337 EST175338

EST175339 EST17534

Contact: HGI (Human Gene Index)

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-0200

Fax: (301)-838-0208

Email: hgi@tigr.org

FEATURES
Location/Qualifiers

source

1..415

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="F0HIBU1"

/lab_host="E. coli DH5-alpha"

/note="Vector: BA, M13-derived; Site_1: HindIII; Site_2:

NotI; The infant brain library, constructed by Bento

Soares, Columbia University, was oligo-(dT) primed and

directionally cloned into an M13-derived plasmid using

total brain mRNA from a 72-day old human female afflicted

with spinal muscular atrophy."

BASE COUNT 82 a 132 c 123 g 71 t 7 others

ORIGIN

Query Match 14.1%; Score 164; DB 10; Length 415;

Best Local Similarity 63.3%; Pred. No. 7.1e-20;

Matches 264; Conservative 0; Mismatches 147; Indels 6; Gaps 1;

Qy 276 accacgtcaacagtgctgcccagctctacattggcgtgacgacgacgctggac 335

Db 5 AACCATCGGAGAGGTCTGGCCAGGCTCTATCTCGGACACGACATGNTAACAC 64

Qy 336 cagctataggctgcagaagcgggggttcacgcacgtgctgaacgcggccacgcgcgtgg 395

Db 65 CCGCGGGAGCTTCGCGCGCTGGGCGATCACGCAGCTCTCAATGCTCACAGCGCGTGG 124

Qy 396 aacgtggacactggccgactactaccgcacatccagatccagatccacgcggtggag 455

Db 125 NCAGGCAC-----GCCGAGGCCCTATGAGGGGNTGGGCATCCGCTACTGGGTGG 178

Qy 456 gccgcagcactgcccacacttcgacactcagtgctctcttccacgcggcgaccttcac 515

Db 179 GCCCAGGACTCGCCAGCCTTTGACATGAGCATCCACTTCACAGCGGTGCCGACTTCATC 238

Qy 516 gacagagcgttaagcagacacacagtaagatcctgttgcactgcgtcatgcccgcagc 575

Db 239 CACCGGGCGGTGAGCCAGCAGGAGGAGATCCTGTGTGCTGTGTGGCGGTGAGC 298

Qy 576 cagtcagcacccttgctcgtcctacctgatgatccacagagcatgaacctggtggac 635

Db 299 CGATCCGCCACCTCGTATGCTGCTACCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 358

Qy 636 gccatccagcagtgcccaagacgcgtgctctctcccgacccgggctttttgaag 692

Db 359 GGCATCAAGAAGTCARAGACCGAGGNATCATCCCAACCGGGGTCTCTGAGG 415

RESULT 14

AK006247

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

AK006247 1065 bp mRNA HTC 05-JUL-2001
Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:1700022L10, full insert sequence.

AK006247 1 GI:12839240

AK006247 CAP trapper.

Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA,
clone:1700022L10.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1065)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Methods in enzymology. 303, 19-44 (1999)

99279253

10349636

2 (bases 1 to 1065)

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome research. 10 (10), 1617-1630 (2000)

20499374

11042159

3 (bases 1 to 1065)

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome research. 10 (11), 1757-1771 (2000)

20530913

11076861

4 (bases 1 to 1065)

The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5 (bases 1 to 1065)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,

Haragaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F.,

Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,

Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T.,

Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,

Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,

Sogabe, Y., Suzuki, H., Shibata, Y., Shinagawa, A., Shiraki, T.,

Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A.,

Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)

Please visit our web site (http://genome.gsc.riken.go.jp/) for

further details.

cDNA library was prepared and sequenced in Mouse Genome


```

Db 209 GCCCAGGGGGCTCTACTGTGAGGCGAGCCCTGACTTCTATGGCAGCAGTGTGAGC--- 265
QY 441 taccacgcgtggagccgacgacctgcccaccttcgacctcagtgcttcttctaccg 500
Db 266 TACCTGGGGGTCCAGCCGACGACCTTCTGAGTTTCGACATCAGTGTCTACTTCTCCTCT 325
QY 501 gggcagccttcacgacagagcgcgtaagcgacgacccacagtaagatcctggtcactgc 560
Db 326 GCAGCTGACTTCATCCACCGTGCCCTTAGCAGCCCTGGGGCCCAAGGTCTCTGGTGCCTGC 385
QY 561 gtcattggcccgagccggtcagccacctgtctctgacctgctgatcccaaggac 620
Db 386 GTGGTGGGGGTGAGCCGCTCCGCCACACTGTGTCTCTGGCTACCTCATGTGCGCCAGCAG 445
QY 621 atgaccctg 629
Db 446 CTCCTCCCTG 454

```

Search completed: February 8, 2002, 19:05:04
 Job time: 3059 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2002, 18:13:10 ; Search time 104.24 Seconds
(without alignments)
9581.580 Million cell updates/sec

Title: US-09-847-519a-1

Perfect score: 1165

Sequence: 1 gccacgtgggggtggctggg.....aaaaaaaaaaaaaaaaaaaa 1165

Scoring table: IDENTITY_NUC

Gapop 10.0 , Capext 1.0

Searched: 930621 seqs, 42862619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_1101.*

- 1: /SID22/qcgdata/geneseq/geneseq/NA1980.DAT.*
- 2: /SID22/qcgdata/geneseq/geneseq/NA1981.DAT.*
- 3: /SID22/qcgdata/geneseq/geneseq/NA1982.DAT.*
- 4: /SID22/qcgdata/geneseq/geneseq/NA1983.DAT.*
- 5: /SID22/qcgdata/geneseq/geneseq/NA1984.DAT.*
- 6: /SID22/qcgdata/geneseq/geneseq/NA1985.DAT.*
- 7: /SID22/qcgdata/geneseq/geneseq/NA1986.DAT.*
- 8: /SID22/qcgdata/geneseq/geneseq/NA1987.DAT.*
- 9: /SID22/qcgdata/geneseq/geneseq/NA1988.DAT.*
- 10: /SID22/qcgdata/geneseq/geneseq/NA1989.DAT.*
- 11: /SID22/qcgdata/geneseq/geneseq/NA1990.DAT.*
- 12: /SID22/qcgdata/geneseq/geneseq/NA1991.DAT.*
- 13: /SID22/qcgdata/geneseq/geneseq/NA1992.DAT.*
- 14: /SID22/qcgdata/geneseq/geneseq/NA1993.DAT.*
- 15: /SID22/qcgdata/geneseq/geneseq/NA1994.DAT.*
- 16: /SID22/qcgdata/geneseq/geneseq/NA1995.DAT.*
- 17: /SID22/qcgdata/geneseq/geneseq/NA1996.DAT.*
- 18: /SID22/qcgdata/geneseq/geneseq/NA1997.DAT.*
- 19: /SID22/qcgdata/geneseq/geneseq/NA1998.DAT.*
- 20: /SID22/qcgdata/geneseq/geneseq/NA1999.DAT.*
- 21: /SID22/qcgdata/geneseq/geneseq/NA2000.DAT.*
- 22: /SID22/qcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1096.6	94.1	1262	22	Human SGP003 phosph
2	1071	91.9	1212	21	DNA encoding a hum
C 3	241	20.7	556	22	Probe #9870 used t
C 4	234	20.1	234	22	Probe #22910 used
5	204.2	17.5	636	22	Human SGP060 phosph
6	204.2	17.5	1878	22	Human polynucleoti
C 7	203.4	17.5	1691	22	Human cDNA SEQ ID
8	202.6	17.4	1491	21	Human cDNA encodin
9	201	17.3	1917	22	Human SGP014 phosph
10	199.4	17.1	904	22	Human phosphatase
11	199.4	17.1	1200	21	Human ORFX ORF2098

12	197.8	17.0	663	21	AAAO9032
13	197.8	17.0	928	21	AAAO9031
14	197.8	17.0	1023	21	AAZ46148
15	197.8	17.0	1300	21	AAC62749
16	196.2	16.8	1380	22	AAI59133
17	175	15.0	1691	22	AAC60226
18	165.2	14.2	597	22	AAF63566
19	161.6	13.9	789	21	AAAO9028
20	161.6	13.9	1016	21	AAAO9027
21	145.8	12.5	4360	22	AAO9494
22	133	11.4	861	15	AAQ68661
23	133	11.4	4580	21	AAC77082
24	123	10.6	451	21	AAC75128
25	119.4	10.2	240	22	AAO9503
26	110	9.4	110	21	AAC16446
27	105.6	9.1	775	21	AAC61101
28	105.6	9.1	1556	22	AAI60298
29	103.4	8.9	7733	22	AAI62831
30	98	8.4	1227	22	AAI59134
31	96.4	8.3	812	22	AAI60919
32	96.4	8.3	812	22	AAI60920
33	70.6	6.1	1238	15	AAQ72864
34	70.6	6.1	1238	16	AAQ56010
35	70.6	6.1	1238	22	AAF68858
36	70.6	6.1	2064	22	AAF68877
37	70.6	6.1	2109	22	AAF68878
38	69	5.9	1619	22	AAF68856
39	69	5.9	4637	22	AAF68859
C 40	68.4	5.9	4580	21	AAC77082
C 41	66.4	5.7	476	22	AAI15696
C 42	66.4	5.7	476	22	AAI37368
C 43	65.2	5.6	230	22	AAI34920
C 44	65.2	5.6	230	22	AAI50550
45	63.2	5.4	35828	21	AAA29063

ALIGNMENTS

RESULT 1

AAD09495

ID AAD09495 standard; DNA; 1262 BP.

XX AAD09495;

XX 10-SEP-2001 (first entry)

XX Human SGP003 phosphatase polypeptide encoding DNA.

Human: SGP003 phosphatase polypeptide; phosphatase-related disease; immune-related disorder; ocular disease; organ transplant rejection; infection; diabetes; pain; sexual dysfunction; Alzheimer's disease; metabolic disorder; haematopoietic cancer; mood disorder; cardiac; Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; cardiovascular disease; brain; neuronal-associated disease; dyskinesia; attention disorder; cognition disorder; psychotic disorder; cystostatic; neurological disorder; viricide; nootropic; cerebroprotective; therapy; neuroprotective; antibacterial; vulnerrary; tranquiliser; antisthmatic; hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive; antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase; MKP; migraine; Chromosome CHR10; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 240..902

XX /*tag= a

XX /product= "Human SGP003 phosphatase polypeptide"

XX /transl_except= (pos:492..494, aa:leu)

XX WO200146394-A2.

XX 28-JUN-2001.

PD

rhumatoid arthritis; microbial infection; trauma; ss.
Homo sapiens.
Key Location/Qualifiers
CDS 172..837
/tag= a
/product= "regulator of intracellular phosphorylation"
W020005332-A2.
21-SEP-2000.
17-MAR-2000; 2000W0-US07277.
18-MAR-1999; 99US-0125593.
20-MAY-1999; 99US-0135049.
09-JUL-1999; 99US-0143188.
XX (INCY-) INCYTE PHARM INC.
XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;
PI Lu DAM, Au-Young J;
XX WPI; 2000-602121/57.
DR P-PSDB; AAB18667.
XX Novel human intracellular phosphorylation regulator polypeptides and
PT polynucleotides for diagnosis, prevention and treatment of
PT neurological, cell proliferative and autoimmune/inflammatory disorders
PT
XX Claim 4; Page 94; 96pp; English.
XX The present sequence encodes a human regulator of intracellular
CC phosphorylation (HRIP). HRIP is useful for screening agonists and
CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
CC are useful for treating a disease or condition associated with
CC decreased or increased expression of functional HRIP. Diseases treated
CC or diagnosed include neurological disorders such as stroke, Parkinson's
CC disease, demyelinating diseases, bacterial and viral meningitis and
CC other developmental disorders of the central nervous system.
CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
CC inflammatory disorder such as Addison's disease, acquired
CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
CC rheumatoid arthritis, microbial infection and trauma.
XX Sequence 1212 BP; 284 A; 345 C; 373 G; 210 T; 0 other;
Query Match 91.9%; Score 1071; DB 21; Length 1212;
Best Local Similarity 97.4%; Pred. No. 4.4e-230;
Matches 1131; Conservative 0; Mismatches 25; Indels 5; Gaps 4;
Qy 2 gccagtgagggtgctggcgctgaggtgctacatgccccacgacagacacaccccccagac 61
Dy 54 gccagtgagggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtg 113
Qy 62 gcggccagggcccgccacacccagctgcagaaaggagagaaatcccttggtcttaaaat 121
Dy 114 gcggccagggcccgccacacccagctgcagaaaggagagaaatcccttggtcttaaaat 173
Qy 122 gacatctggagaagtgaagacacccctcaagaatgcctactatctgcacagaggtgtc 181
Dy 174 gacatctggagaagtgaagacacccctcaagaatgcctactatctgcacagaggtgtc 233
Qy 182 gccagaatgag 241
Dy 234 gccagaatgag 293

RESULT 3

AAI41184/C

ID AAI41184 standard; DNA; 556 BP.

XX AAI41184;

XX AC AAI41184;

XX DT 17-OCT-2001 (first entry)

XX

erythematous, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, thrombotic thrombolytic disease; to enhance coagulation; to inhibit thrombolysis; and as a contraceptive.

Query Match	17.1%;	Score 199.4;	DB 21;	Length 1200;
Best Local Similarity	64.3%;	Pred. NO. 2.1e-35;		
Matches 299; Conservative	0;	Mismatches 166;	Indels 0;	Gaps 0;

Qy	276	accacgcgtcaacagagctctggcccagcctacatctggcgtacgagcgcgctggac	335
Db	490	aaccatcgtatgagctctggcccagcctctctcctggagatgctacgagcccgggac	549
Qy	336	cgctatagctgcagaagcgcggttccagcacgctgctaacgcgcccacggccgctgg	395
Db	550	aagagcaagctgatccagctgggaaacaccacgttgtgaatgcgcgtgcagggcaagtcc	609
Qy	396	aacgtgcacactgggcccgcactactaccgcgacatggacatccagtacacgcgctggag	455
Db	610	caggtggacacaggtgccaaattccactcggaatgtccctggagtactatgcgcatcgag	669
Qy	456	gccgacgacctgcacaccttcgcactcagtgctctctctaccccggcgacgccttcac	515
Db	670	gcggagacacacccctctctgcacctcagtgctactcttctgctgtgtctcgatcac	729
Qy	516	gacagagcgttaagcagcacaccacagtaagatcctggttcactgcgctcatgagccgcagc	575
Db	730	cgagctccctcagtggtccccaagcgccgctgctggtacactgtgccatggggtaagc	789
Qy	576	cggtcagccacctggtctctggcctacctgatgatccacaagacatgacctggtggac	635
Db	790	cgctctgccacctgtctctggccttcctcatgatctatgagaacatgacgtggttagag	849
Qy	636	gccatccagcaagtggccaagaccgctcgctcctccgaacccgggctttttgaagcag	695
Db	850	gccatccagcaggtgtagggcccaccgcgaatatctgccttaactcagaggtctctccggcag	909
Qy	696	ctccgggagctgcacagcagctggtgcgacgagggcgacggtcc	740
Db	910	ctccagttctgacacaccacactggagcgaggagacacggcgagcttc	954

RESULT 12

AAA09032
 ID AAA09032 standard; DNA; 663 bp.
 XX
 XX
 AC AAA09032;
 XX
 DT 01-AUG-2000 (first entry)
 XX
 DE Human CSATP-4 open reading frame.

Cardiovascular system associated protein tyrosine phosphatase 4;
CSA TP-4; cytostatic; immunomodulatory; antidiabetic; virucide;
hypotensive; cardiant; tyrosine phosphatase modulator; ss.

OS Homo sapiens.

XX
PN
WO200018890-A2.XX
PD
06-APR-2000XX
PF
30-SEP-1999: 99W0-IIS22924XX
30-SEP-1998. 98US-0163833

PR 30-SEP-1998; 98US-0164193.

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Acton S;
XX
DR WPI; 2000-293136/25
DR P-PSDB; AAY92177.

New cardiovascular system associated protein tyrosine phosphatases are used in the diagnosis and treatment of e.g. immune disorders, anti-proliferative disorders, metabolic disorders and cardiovascular disorders

PS Claim 1; Page 155-156; 156pp; English.

This sequence encodes a cardiovascular system associated protein tyrosine phosphatase 4 (CSA4PTP). The CSA4PTP nucleic acid and protein molecules are used to modulate regulation of cellular processes. CSA4PTP nucleic acid and protein molecules and modulators of CSA4PTP activity and expression can be used to treat a subject with a disorder characterized by aberrant CSA4PTP expression or activity. These disorders can include an immune disorder, an anti-proliferative disorder, a proliferative disorder e.g. renal and lung carcinomas, a metabolic disorder e.g. diabetes, viral pathogenesis, a neural disorder, a cardiovascular disorder e.g. hypertension or coronary heart disease or a disorder arising from improper phosphorylation of a phosphorylated protein. Anti-CSA4PTP antibodies are used to isolate CSA4PTP by standard techniques, to facilitate the purification of natural and recombinantly produced CSA4PTP from cells, to detect CSA4PTP protein in cell supernatant for evaluating the level of CSA4PTP expression and to monitor protein levels in tissue as part of a clinical testing procedure to determine efficacy of a treatment procedure. CSA4PTP nucleic acids, antibodies and protein molecules can be used in screening assays and in predictive medicine e.g. prognostic assays, monitoring clinical trials and pharmacogenetics.

SQ Sequence 663 BP; 126 A; 222 C; 182 G; 133 T; 0 other;

	Query Match	Score	DB	Length
Best Local Similarity	64.1%	197.8	21	663
Matches	298	Pred. No. 4e-35		
Conservative	0	Mismatches	167	Indels
				Gaps
				0

Qy	276	accacgctcaacgaggtctctggcccaagctctacattggcgatgagcgacgctgcgtggac	335
Db	199	aacatactgatgaggtctggcccgacctctctctggagatgctacgacgcccggac	258
Qy	336	cgctataggctgcagaagcgcggggttcaagcacgtgctgaacgcggccacagggcgcgtgg	395
Db	259	aagagaagctgtaccagctgggaatacaccacgctgtgaaatgcgcctgcaggcaagtttc	318
Qy	396	aacgtggacactgggcccgaactactaccgcgcagatccagatccagataccacagcgctggag	455
Db	319	caggtggacacaggtgccaaattctaccgtggaaatgccctggagtactatggcattgag	378
Qy	456	gcgcagcagactgcccaacttcgacctcagtgctctcttctaccgcggcagccttcctc	515
Db	379	gcgggaacaaacccctctctcgacctcagtgctactctctgctgttctcgatatac	438
Qy	516	gacagagcgttaagcgacgacaccacagtaagatcctggttcaactgcgtctatggccgcagc	575
Db	439	cgagctgcctcagtgttcccaagccgcgctgctagtacactgtgccatgggggtaagc	498
Qy	576	cggttcagccacctggtctctggcctaactgatgatccaaagaagacatgaccttggtggac	635
Db	499	cgctctgccacactgtctctggcctctctcatgatctatgagaacatgacgtggttagag	558
Qy	636	gccatccagcaagtgggccaaagaccgctggctcctccgaacccgggcgtttttgaagcag	695
Db	559	gcaatccagacggttgaggcccaaccgaatactgcccctaactcaggettctccggcag	618
Qy	696	ctccgggagctgcacaagcagctggttgcagcagagggcagcggtcc	740
Db	619	ctccagttctgacaacacagctggggcgagacacacggcggttc	663

RESULT 13

Oy 696 ctccgggagctggacaagcagctgtgcagcagagcgacggtcc 740
||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1010 ctccaggttctggacaaccgactgggcccggagagacggggcggttc 1054

Search completed: February 8, 2002, 18:45:16
Job time: 1926 sec

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:41:46 : Search time 23.44 Seconds
(without alignments)
1372.863 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 1161

Sequence: 1 MTSGEVKTSLKNAYSAKRL.....VOQRRSRQDGEEDGREL 220

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	461	39.7	198	11 Q9QXJ7	Q9QXJ7 mus musculus
2	460.5	39.7	198	4 Q9UII6	Q9UII6 homo sapien
3	455	39.2	198	11 Q9DA25	Q9DA25 mus musculus
4	445	38.3	211	4 Q9BV47	Q9BV47 homo sapien
5	440	37.9	198	11 Q9D700	Q9D700 mus musculus
6	345.5	29.8	245	5 Q9VWN2	Q9VWN2 drosophila
7	335	28.9	185	11 Q9D7X3	Q9D7X3 mus musculus
8	233	20.1	86	4 Q9BTW0	Q9BTW0 homo sapien
9	205	17.7	184	4 Q9NRW4	Q9NRW4 homo sapien
10	204.5	17.6	365	5 Q4L128	Q4L128 caenorhabdi
11	203.5	17.5	198	10 Q9ER37	Q9ER37 arabisopsis
12	203.5	17.5	198	10 Q9LUG6	Q9LUG6 arabisopsis
13	201.5	17.4	946	10 Q9FFA8	Q9FFA8 arabisopsis
14	200	17.2	184	11 Q9GN11	Q9GN11 mus musculus
15	199.5	17.2	167	10 Q9MBK7	Q9MBK7 arabisopsis
16	193	16.6	436	11 Q9KAC2	Q9KAC2 mus musculus
17	193	16.6	1045	5 Q9NKL1	Q9NKL1 drosophila
18	192.5	16.6	303	4 Q9NSW1	Q9NSW1 homo sapien
19	192.5	16.6	394	4 Q13524	Q13524 homo sapien

20	192.5	16.6	411	4 Q13649	Q13649 homo sapien
21	190	16.4	661	10 Q9ATY4	Q9ATY4 zea mays (m
22	185	15.9	784	10 Q9C5S1	Q9C5S1 arabisopsis
23	182	15.7	608	5 Q9VU80	Q9VU80 drosophila
24	181.5	15.6	369	13 Q91790	Q91790 xenopus lae
25	179	15.4	665	4 Q9BY84	Q9BY84 homo sapien
26	179	15.4	690	4 Q9C0G3	Q9C0G3 homo sapien
27	177.5	15.3	198	4 Q95147	Q95147 homo sapien
28	176.5	15.2	353	13 Q42253	Q42253 gallus gall
29	174.5	15.0	482	4 Q9Y6W6	Q9Y6W6 homo sapien
30	174	15.0	220	11 Q9D6P6	Q9D6P6 mus musculus
31	174	15.0	220	11 Q99N12	Q99N12 mus musculus
32	174	15.0	223	11 Q60970	Q60970 mus musculus
33	174	15.0	223	11 Q9DCF8	Q9DCF8 mus musculus
34	173.5	14.9	476	5 Q9VHV8	Q9VHV8 drosophila
35	173.5	14.9	483	11 Q9CZY9	Q9CZY9 mus musculus
36	173	14.9	677	11 Q99NG6	Q99NG6 mus musculus
37	172.5	14.9	177	11 Q9CSL5	Q9CSL5 mus musculus
38	172.5	14.9	198	11 Q9D715	Q9D715 mus musculus
39	172.5	14.9	483	11 Q9ESS0	Q9ESS0 mus musculus
40	170.5	14.7	476	5 Q46122	Q46122 drosophila
41	169.5	14.6	189	11 Q9D9D8	Q9D9D8 mus musculus
42	169.5	14.6	198	11 Q9JLY7	Q9JLY7 mus musculus
43	169	14.6	348	5 Q9VVM4	Q9VVM4 drosophila
44	168	14.5	205	11 Q60969	Q60969 mus musculus
45	168	14.5	226	5 Q93592	Q93592 caenorhabdi

ALIGNMENTS

RESULT 1					
Q9QXJ7					
ID	Q9QXJ7	PRELIMINARY;	PRT;	198 AA.	
AC	Q9QXJ7;				
DT	01-MAY-2000 (TrEMBLrel. 13, Created)				
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)				
DE	PROTEIN PHOSPHATASE (DUAL-SPECIFICITY PHOSPHATASE TS-DSP6).				
GN	DUSP13.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20053896; PubMed=10585869;				
RA	Nakamura K., Shima H., Watanabe M., Haneji T., Kikuchi K.;				
RT	"Molecular cloning and characterization of a novel dual-specificity protein phosphatase possibly involved in spermatogenesis.";				
RL	Biochem. J. 344:819-825(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Aoyama K., Matsuda T., Aoki N.;				
RT	"Molecular cloning of a novel dual specificity phosphatase.";				
KL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF027003; BAA89411.1; -				
DR	EMBL; AF237620; AAK15037.1; -				
DR	HSSP; P51452; 1VHR.				
DR	MGD; MGI:1351599; Dusp13.				
DR	InterPro; IPR000340; DS_phosphatase.				
DR	InterPro; IPR000387; TYR_phosphatase.				
DR	Pfam; PF00782; DSPC; 1.				
DR	SMART; SM00195; DSPC; 1.				
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.				
DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.				
DR	PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.				
KW	Hydrolase.				
SQ	SEQUENCE 198 AA; 22481 MW; C2493597D6D3732B CRC64;				

Query Match 39.7%; Score 461; DB 11; Length 198;
Best Local Similarity 49.7%; Pred. No. 5.4e-36.


```

DR EMBL; AK009781; BAB26501.1; -
DR MGDB; MGI:1914209; 2310043K02Rik.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
SQ SEQUENCE 198 AA; 22419 MW; 11F0CAF4B5620F0E CRC64;

```

[illegible]

01-MAY-2000	(TEMBLurel. 13, Last sequence update)
01-JUN-2001	(TEMBLurel. 13, Last sequence update)
01-JUN-2001	(TEMBLurel. 17, Last annotation update)
CG37378	PROTEIN.
CG37378	
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephygroidae; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
OX	[1]
RP	SEQUENCE FROM N.A.
RP	STRAIN=BERKELEY;
RC	MDLNAME=20196006; PubMed=10731132;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA	Ballem R.M., Basu A., Bayendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hosdin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Moberg C., Morris J., Moshrefi A.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinch S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RL EMBL, AK008734; BAB25864.1; ...
DR MGD; MG1:1919599; 2210015003Rik.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM001195; DSPC; 1.
DR SMART; SM000112; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE; PS00506; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00504; TYR_PHOSPHATASE_DUAL; 1.
DR PROSITE; 185 AA; 20472 MW; 62E519E41BE575D5 CRC64;
SQ

[illegible]

		Query Match	20.1%; Score 233; DB 4;	Length 86;	
		Best Local Similarity	58.5%;	Pred. No.	7 6e-15;
		Matches	48;	Conservative	15; Mismatches 19; Indels 0; Gaps
Dt	JUN-2001 (JEMBLrel. 17,				
Tl	-JUN-2001 (TJEBMLrel. 17,				
Oc	UNKNOWN (PROTEIN FOR MGC:2627).				
Sr	Homo sapiens (Human)				
Cx	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Xm	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
Nb	NCBI_TaxID=9606;				
Rn	[1]				
Rp	SEQUENCE FROM N.A.				
Rc	TISSUE=NEUROBlastoma;				
Rl	Straussberg R.;				
Rd	Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.				
Dr	EMBL: BC003115; AAH03115.1;				
Sq	SEQUENCE 86 aa; 9652 MW; 4DD29AFD5989528B CRC64;				
Qy	121 LSVEFFYPAAAFIDRALSDSHSKILVHCVMGRSRSATLVLAYLMHKMDTLVDIAQQAKN				180
	: : :				:
Db	1 MSIFQTAADEFIHRALSOPGKILVHCVCVSRSATLVLAYLMYHLTVLEAIKKVKDH				60
Qy	181 RCVLPNRGFLQLRDLKDQLVV				202
	: :				:
Db	61 RGIIPINRGFULLALDRRLRO				82

RESULT 9

Q9NRW4 ID Q9NRW4 PRELIMINARY; PRT; 184 AA.
AC Q9NRW4
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE X.
GN MKPX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gu J., Huang O., Yu Y., Xu S., Wang Y., Han Z., Chen Z., Zhou J.,
Tu Y., Gu W., Fu G., Huang C.;
RA "Novel genes expressed in hematopoietic stem/progenitor cells from
RT myelodysplastic Syndromes patient."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF165519; AAF66649.1;
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Kinase.
SQ SEQUENCE 184 AA; 20910 MW; B3F962A087C2BA20 CRC64;

Query Match 17.7%; Score 205; DB 4; Length 184;
Best Local Similarity 35.7%; Pred. No. 9.3e-12;
Matches 61; Conservative 22; Mismatches 66; Indels 22; Gaps 5;

QY 55 VNEVWPKLYIGDEATDRLRYLQKAGFTHLVNAHGRNVDTGPDYRDM--DIQVHGVE 112
DB 5 MNKILPGLYIGNFKDARDAEQLSKNKKVTHLSV-----HDSRPMLEGVKYLCP 54
QY 113 ADDLFTFDSLVFFYPAAAFIDRALSDHSHKILVHCVMGRSATSILVLAHMTKDMTLVD 172
DB 55 AADSPSQNLTRHFESIKFIHEICRLRGES-CLVHCLAGVSRVTLVIAVIMTVDFGWED 113
QY 173 AIQOVAKNR-CVLPNGFLKQLRELDKQLVQORR-----SQRODGE 214
DB 114 ALHTVRAGRSCANPNVGRQLOEFKEHGVQYRWLKEEYGESPLQDAE 164

RESULT 10

O44128 ID O44128 PRELIMINARY; PRT; 365 AA.
AC O44128
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE C05B10.1 PROTEIN.
GN C05B10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Geisel C., Wamsley P.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF036685; AAB88308.1;
DR HSSP; Q16828; IMKP.
DR InterPro; IPR001763; Rhodanese domain.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 365 AA; 41101 MW; 1E416C0E9693AF66 CRC64;

Query Match

Best Local Similarity 17.6%; Score 204.5; DB 5; Length 365;
Matches 54; Conservative 24; Mismatches 55; Indels 7; Gaps 3;

QY 62 LVIGDEATDRLRYLQKAGFTHLVNAHGRNVDTGPDYRDMIOYHGVEADDLFTFDL 121
DB 188 LYLGAETAKNRDLVKLKKYSISHVINTSLNPT-----FEEDPNMYRLRISADDNASHNL 242
QY 122 SVFFYPAAAFIDRALSDHSHKILVHCVMGRSATSILVLAHMTKDMTLVDAIQOVAK-N 180
DB 243 TKFFPEAFSFDIARRND-SACLVHCLAGISRSVTLCLAYLMKTEMCTLDSAYENVQKRKN 301
QY 181 RCVLNPGFLKQLRELDKQL 200
DB 302 ASIAPNFHEMGOLTDYEKML 321

RESULT 11
Q9ZR37 ID Q9ZR37 PRELIMINARY; PRT; 198 AA.
AC Q9ZR37
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE DSPTP1 PROTEIN.
GN DSPTP1
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99155801; PubMed=10036776;
RA Gupta R., Huang Y., Kieber J., Luan S.;
RT "Identification of a dual-specificity protein phosphatase that
RT inactivates a MAP kinase from Arabidopsis";
RL Plant J. 16:581-589(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Gupta R.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y18620; CAA77232.1;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:44:11 ; Search time 12.75 seconds
(without alignments)
388.292 Million cell updates/sec

Title: US-09-847-519A-2
Perfect score: 220
Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VQORRRSQRDGEEDGREL 220

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 212352 seqs, 22503292 residues

Word size : 10

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	12	5.5	207	4	US-09-013-881-4
2	12	5.5	263	4	US-09-164-193-5

ALIGNMENTS

RESULT 1

US-09-013-881-4
; Sequence 4, Application US/09013881
; Patent No. 6132964

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Shah, Purvi

TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013.881
FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: BILLINGS, LUCY J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0470 US

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 207 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: MUSCNOT02

CLONE: 971204

US-09-013-881-4

Query Match 5.5%; Score 12; DB 4; Length 207;

Best Local Similarity 100.0%; Pred. No. 0.00039;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 SRSATLVLAYLM 163

Db 153 SRSATLVLAYLM 164

|||||

RESULT 2

US-09-164-193-5

; Sequence 5, Application US/09164193C

; Patent No. 6258582

GENERAL INFORMATION:

APPLICANT: Acton, Susan L.

TITLE OF INVENTION: NOVEL CSAPTP NUCLEIC ACID MOLECULES AND USES THEREFOR (as ame

FILE REFERENCE: MNI-051

CURRENT APPLICATION NUMBER: US/09/164,193C

CURRENT FILING DATE: 1998-09-30

NUMBER OF SEQ ID NOS: 23

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5

LENGTH: 263

TYPE: PRT

ORGANISM: Homo sapiens

US-09-164-193-5

Query Match 5.5%; Score 12; DB 4; Length 263;

Best Local Similarity 100.0%; Pred. No. 0.00048;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 SRSATLVLAYLM 163

Db 114 SRSATLVLAYLM 125

|||||

Search completed: February 8, 2002, 15:46:05

Job time: 114 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 8, 2002, 15:42:06 ; Search time 11.67 seconds
(without alignments)
691.196 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 1161

Sequence: 1 MTSGEYKTSKLNAYSAKRL.....VOQRRRSQRQDEEDGREL 220

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	346.5	29.8	185	1	DUS3_HUMAN	P51452	homo sapien
2	196.5	16.9	384	1	DUS9_HUMAN	Q99956	homo sapien
3	192.5	16.6	394	1	DUS4_HUMAN	Q13115	homo sapien
4	192.5	16.6	395	1	DUS4_RAT	Q62767	rattus norv
5	184.5	15.9	318	1	DUS2_MOUSE	Q05922	mus musculus
6	180.5	15.5	314	1	DUS2_HUMAN	Q05923	homo sapien
7	180	15.5	375	1	DUS4_CHICK	Q9PW71	gallus gall
8	169	14.6	367	1	DUS1_HUMAN	P28562	homo sapien
9	168.5	14.5	384	1	DUS5_RAT	O54838	rattus norv
10	166	14.3	280	1	DUS7_RAT	Q63340	rattus norv
11	166	14.3	322	1	DUS7_HUMAN	Q16829	homo sapien
12	166	14.3	367	1	DUS1_MOUSE	P28563	mus musculus
13	166	14.3	367	1	DUS1_RAT	Q64623	rattus norv
14	163	14.0	625	1	DUS8_HUMAN	Q13202	homo sapien
15	163	14.0	663	1	DUS8_MOUSE	O09112	mus musculus
16	161.5	13.9	381	1	DUS6_HUMAN	Q16828	homo sapien
17	161.5	13.9	381	1	DUS6_RAT	Q64346	rattus norv
18	157.5	13.6	384	1	DUS5_HUMAN	Q16690	homo sapien
19	136.5	11.8	276	1	PTP3_CHLEU	Q39491	chlamydomon
20	132.5	11.4	619	1	VHPI1_CAEEL	Q10038	caenorhabdi
21	132	11.4	209	1	YIL3_YEAST	P40479	saccharomyc
22	128.5	11.1	340	1	DUSC_HUMAN	Q9uni6	homo sapien
23	124.5	10.7	489	1	MSG5_YEAST	P38590	saccharomyc
24	122	10.5	364	1	PVH1_YEAST	Q02256	saccharomyc
25	113	9.7	171	1	VH01_VACCV	P07239	vaccinia vi
26	111	9.6	292	1	Y042_CAEEL	P34680	caenorhabdi
27	108	9.3	171	1	VH01_RACVI	P80994	raccoon pox
28	107	9.2	171	1	VH01_VACCV	P20495	vaccinia vi
29	106	9.1	171	1	VH01_VARV	P33064	variola vir
30	98	8.4	160	1	PTP2_NPVOF	O10273	orgyia pseu
31	97.5	8.4	1442	1	PTPG_MOUSE	Q05909	mus musculus
32	93.5	8.1	1445	1	PTPG_HUMAN	P23470	homo sapien
33	91	7.8	551	1	CC14_YEAST	Q00684	saccharomyc

RESULT 1

ID	DUS3_HUMAN	STANDARD;	PRT;	185 AA.
AC	P51452;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	DUAL SPECIFICITY PROTEIN PHOSPHATASE 3 (EC 3.1.3.48) (EC 3.1.3.16)			
DE	(DUAL SPECIFICITY PROTEIN PHOSPHATASE VHR).			
GN	DUSP3 OR VHR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RA	MEDLINE=93101689; PubMed=1281549;			
RA	Ishibashi T., Bottaro D.P., Chan A., Miki T., Aaronson S.A.;			
RT	"Expression cloning of a human dual-specificity phosphatase.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 89:12170-12174(1992).			
RN	[2]			
RP	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).			
RX	MEDLINE=96243129; PubMed=8650341;			
RA	Yuvaniyama J., Denu J.M., Dixon J.E., Saper M.A.;			
RT	"Crystal structure of the dual specificity protein phosphatase VHR.";			
RL	Science 272:1328-1331(1996).			
CC	!- FUNCTION: THIS PROTEIN SHOW BOTH ACTIVITY TOWARD TYROSINE-			
CC	PROTEIN PHOSPHATE AS WELL AS WITH SERINE-PROTEIN PHOSPHATE.			
CC	!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O -			
CC	PROTEIN TYROSINE + ORTHOPHOSPHATE.			
CC	!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-			
CC	TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L05147; AAA35777.1;			
DR	PDB; 1VHR; 20-JUN-96.			
DR	MIM; 600183;			
DR	InterPro; IPR000340; DS_phosphatase.			
DR	InterPro; IPR000387; TYR_phosphatase.			
DR	Pfam; PF00782; DSPC; 1.			
DR	SMART; SM00195; DSPC; 1.			
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.			
DR	PROSITE; PS0056; TYR_PHOSPHATASE_2; 1.			
DR	PROSITE; PS0054; TYR_PHOSPHATASE_DUAL; 1.			
KW	Hydrolase; 3D-structure.			
FT	ACT_SITE 124 124			
SQ	SEQUENCE 185 AA; 20478 MW; C1045DD9B226FD94 CRC64;			

Q98936 gallus gall
P38148 saccharomyc
P22105 homo sapien
Q62656 rattus norv
P48441 mus musculus
P76093 escherichia
P54252 homo sapien
P23471 homo sapien
P43078 candida alb
Q941k5 drosophila
P70569 rattus norv
P20417 rattus norv

ALIGNMENTS

Query Match 29.8%; Score 346.5; DB 1; Length 185;
Best Local Similarity 43.8%; Pred. No. 8.2e-27;
Matches 78; Conservative 35; Mismatches 54; Indels 11; Gaps 3;

QY 37 GAFE-----LERLFKSGPQT-----HYNEVWPKLYIGDEATDRLQKAGTHVLNA 87
DQ 3 GSFELSDVLDLSDGSGCYSLPQPCNEVTPRIYVGNASVAQDIPKQLKIGITHVLNA 62
QY 88 AHGR--WVVDYGDYRDMDIOYHGVADDLPTDLSVEFFYPAAAFIDRLSDHSHKTLV 145
DQ 63 AGRFHMVNTNANIKSGYILGKANDYQEFNLSAYFAERADFDQALAKNGRVLV 122
QY 146 HCVGMRSSATLVLAYLMIHKDMTLDVDAIQVAKNRCVLPNRGFLKQLRELDKQLVQ 203
DQ 123 HCRCYRSPTLVIAYLMWRQKMDKVSALSIVRQNRREIGPNDGFLAQLCOLNDRLE 180

RESULT 2
ID DUS9_HUMAN STANDARD; PRT; 384 AA.
AC Q99556;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 9 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE 4) (MAP KINASE
DE PHOSPHATASE 4) (MKP-4).
GN DUSP9 OR MKP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA MEDLINE=97184169; PubMed=9030581;
RA Muda M., Boschart U., Smith A., Antonsson B., Gillieron C.,
RT Chabert C., Camps M., Martinou I., Ashworth A., Arkinstall S.;
RT "Molecular cloning and functional characterization of a novel
RT mitogen-activated protein kinase phosphatase, MKP-4";
RL J. Biol. Chem. 272:5141-5151(1997).
CC -!- FUNCTION: INACTIVATES MAP KINASES. HAS A SPECIFICITY FOR THE ERK
CC FAMILY.
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATASE + H(2)O =
CC PROTEIN TYROSINE + PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; Y08302; CAA69610.1; -
DR MTM; 300134; -
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
FT DOMAIN 203 384 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 290 290 BY SIMILARITY.
FT SEQUENCE 384 AA; 41867 MW; F8598CA95AB379B7 CRC64;

Query Match 16.9%; Score 196.5; DB 1; Length 384;
Best Local Similarity 30.5%; Pred. No. 9.7e-12;
Matches 62; Conservative 34; Mismatches 82; Indels 25; Gaps 7;

QY 12 NAYSARK--LSPKMEEGEEDYCTPGAFELERLFKSGPQYTHVNEVWPKLYIGDEAT 69
DQ 173 DAEEADRDMSGCLDSEG-----ATPPVGLRASFD-----PVQILPNLYLSARD 218
QY 70 ALDRYRLOKAGFTHVNLNAHGRWNVDTGPDY--RDMDTQYHGVREADDLPTDLSVFFYPA 128
DQ 219 SANLESALAKGIRYLINVTPLN-----PNFEKNGDPHYKQIPISDHSQNSLSRFFPEA 272
QY 129 AAFIDRLASDDHSHKTLVHCVGMRSSATLVLAYLMIHKDMTLDVDAIQVAKNRC-VLPNR 187
DQ 273 IEFIDEALS-QNCGLVHCLAGSRSVTVTVAYLMQKHLHLSLNDAYDLVKKRKSINSPNF 331
QY 188 GFLKQLRELDKQLVQRRRSORQ 210
DQ 332 NFMGQLDFFERSLREERHSOEQ 354

RESULT 3
ID DUS4_HUMAN STANDARD; PRT; 394 AA.
AC Q13115;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE HVH2).
GN DUSP4 OR VH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95221370; PubMed=7535768;
RA Guan K.-L., Butch E.;
RT "Isolation and characterization of a novel dual specific phosphatase,
RT HVH2, which selectively dephosphorylates the mitogen-activated
RT protein kinase";
RL J. Biol. Chem. 270:7197-7203(1995).
CC -!- FUNCTION: REGULATES MITOGENIC SIGNAL TRANSDUCTION BY
CC DEPHOSPHORYLATING BOTH THR AND TYR RESIDUES ON MAP KINASES
CC ERK1 AND ERK2.
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATASE + H(2)O =
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U21108; AAA85119.1; -
DR HSP; P51452; 1VHR.
DR MTM; 602747; -
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.

```

DR EMBL: U23438; AAC52493.1; -
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR001763; Rhodanese_domain.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR SMART: SM00519; Rhodanese; 1.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS00556; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS0054; TYR_PHOSPHATASE_DUAL; 1.
DR KW Hydroxylase; Nuclear protein.
FT DOMAIN 47 62 CH2 A DOMAIN.
FT DOMAIN 139 154 CH2 B DOMAIN.
FT DOMAIN 198 395 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 281 281 BY SIMILARITY.
SQ SEQUENCE 395 AA; 43187 MW; A908FFFD378A050FD CRC64;

```

[illegible]

RESULT	5
DUS2_MOUSE	
ID	DUS2_MOUSE STANDARD; PRT; 318 AA.
AC	Q05922; Q60840;
DT	01-FEB-1994 (Rel. 28, Created)
DT	01-FEB-1994 (Rel. 28, Last sequence update)
DE	01-NOV-1997 (Rel. 35, Last annotation update)
DE	DUAL SPECIFICITY PROTEIN PHOSPHATASE 2 (EC 3.1.3.48) (EC 3.1.3.16)
DE	(DUAL SPECIFICITY PROTEIN PHOSPHATASE PAC-1).
GN	DUSP2 OR PAC1 OR PAC-1.
DN	Mus musculus (Mouse).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE-93206122; PubMed-7681221;
RA	Rohan P., Davis P., Moskaluk C.A., Kearns M., Kruttsch H.,
RA	Siebenlist U., Kelly K.;
RT	"PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.";
RL	Science 259:1763-1766(1993).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=J29;
RX	MEDLINE-95203877; PubMed-7896276;
RA	Gerondakis S., Economou C., Grumont R.J.;
RT	"Structure of the gene encoding the murine dual specificity tyrosine-
RT	threonine phosphatase PAC1."
RL	Genomics 24:182-184(1994).
CC	-1- FUNCTION: REGULATES MITOGENIC SIGNAL TRANSDUCTION BY
CC	DEPHOSPHORYLATING BOTH THR AND TYR RESIDUES ON MAP KINASES
CC	ERK1 AND ERK2.
CC	-1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O -
CC	PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.
CC	-1- TISSUE SPECIFICITY: IN HEMATOPOIETIC TISSUES SUCH AS SPLEEN AND
CC	THYMUS.

CC -!- INDUCTION: BY MITOGENS.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L11330; AAA19566.1; -
CC DR EMBL: U09268; AA85136.1; -
CC DR HSP: P51452; IVHR.
CC MGD: MGI:101911; Dusp2.
CC DR InterPro: IPR000340; DS_phosphatase.
CC DR InterPro: IPR001763; Rhodanese_domain.
CC DR InterPro: IPR000387; TYR_phosphatase.
CC DR Pfam: PF00782; DSPC; 1.
CC DR Pfam: PF00581; Rhodanese; 1.
CC DR SMART: SM00195; DSPC; 1.
CC DR SMART: SM00450; RHOD; 1.
CC DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
CC DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
CC DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
CC KW Hydrolase; Nuclear protein; Alternative splicing.
CC FT ACT_SITE 261 BY SIMILARITY.
CC FT VARSPIC 175 GGPVE -> VSSDL (IN SHORT ISOFORM).
CC FT CONFLICT 180 318 MISSING (IN SHORT ISOFORM).
CC FT CONFLICT 11 12 CE -> WQ (IN REF. 2).
CC FT CONFLICT 20 20 A -> V (IN REF. 2).
CC FT CONFLICT 156 156 P -> A (IN REF. 2).
CC SQ SEQUENCE 318 AA; 34546 MW; A2006ED0FD27E41D CRC64;

Query Match 15.9%; Score 184.5; DB 1; Length 318;
Best Local Similarity 28.3%; Pred. No. 1.1e-10;
Matches 56; Conservative 39; Mismatches 84; Indels 19; Gaps 6;
QY 10 LKNAYSSAKRLSPKMEEGEDYCTPGAFELER-----LFWKGSPTQTHVNEWPKLYI 64
DB 132 LRGGFKSFQYCPDLCSEAPQAQALPPAGAENSNDPRVPYDQGEV-----EILPYLYL 186
QY 65 GDEATALDRYLOKAGFTHVLAHGRWNVDTGPDYRDMDIQYHGVEADLPTFDLSVF 124
DB 187 GSCNISSDLQGLQACGITAFLNV-----ASCNHEEGL-FHYKSIPIVEDNQWVEISAW 239
QY 125 FYPAAFIDRALSDHRSKILVHCVMGRSRSATLVLAHGRWNVDTGPDYRDMDIQYHGVEADLPTFDLSVF 184
DB 240 FQEAISFID-SVKNSSGGRVHLVHCQAGISRSATICLAYLIQSHRVRLEDAFDFVKORRGVI 298
QY 185 -PNRGFLKQLRELDKQLV 201
DB 299 SPNFSFMGLQLQLETVL 316
RESULT 6
DUS2_HUMAN
ID DUS2_HUMAN STANDARD; PRT; 314 AA.
AC Q05923;
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 2 (EC 3.1.3.48) (EC 3.1.3.16)
DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE PAC-1).
GN DUSP2 OR PAC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=93206122; PubMed=7681221;
RA Rohan P., Davis P., Moskaluk C.A., Kearns M., Krutzsch H.,
RA Siebenlist U., Kelly K.;
RT "PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.";
RL Science 259:1763-1766(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96070437; PubMed=7590752;
RA Yi H., Morton C.C., Weremowicz S., McBride O.W., Kelly K.;
RA "Genomic organization and chromosomal localization of the DUSP2 gene,
RT encoding a MAP kinase phosphatase, to human 2p11.2-q11.";
RL Genomics 28:92-96(1995).
CC -!- FUNCTION: REGULATES MITOGENIC SIGNAL TRANSDUCTION BY
CC DEPHOSPHORYLATING BOTH THR AND TYR RESIDUES ON MAP KINASES
CC ERK1 AND ERK2.
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATASE + H(2)O =
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- TISSUE SPECIFICITY: IN HEMATOPOIETIC TISSUES.
CC -!- INDUCTION: BY MITOGENS.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L11329; AAA50779.1; -
CC DR EMBL: U23853; AA86112.1; -
CC DR HSP: P51452; IVHR.
CC MIM: 603068; -
CC DR InterPro: IPR000340; DS_phosphatase.
CC DR InterPro: IPR001763; Rhodanese_domain.
CC DR InterPro: IPR000387; TYR_phosphatase.
CC DR Pfam: PF00782; DSPC; 1.
CC DR Pfam: PF00581; Rhodanese; 1.
CC DR SMART: SM00195; DSPC; 1.
CC DR SMART: SM00450; RHOD; 1.
CC DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
CC DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
CC DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
CC KW Hydrolase; Nuclear protein.
CC FT ACT_SITE 257 257 BY SIMILARITY.
CC SQ SEQUENCE 314 AA; 34399 MW; FDD3543C6DE10CA5 CRC64;
Query Match 15.5%; Score 180.5; DB 1; Length 314;
Best Local Similarity 32.2%; Pred. No. 2.7e-10;
Matches 47; Conservative 32; Mismatches 58; Indels 9; Gaps 4;
QY 57 EWPKLYICDEATADLRYLKAGFTHVLAHGRWNVDTGPDYRDMDIQYHGVEADL 116
DB 175 EILPYFLGSCSHSSDLQGLQACGITAFLNV-----ASCNHEEGL-FRYKSIPIVEDN 227
QY 117 PTFDLSVFFYPAAAFIDRALSDHRSKILVHCVMGRSRSATLVLAHGRWNVDTGPDYRDMDIQYHGVEADL 176
DB 228 QMVEISAWFQEAIGFID-WFKNSGGRVHLVHCQAGISRSATICLAYLIQSHRVRLEDAFDF 286
QY 177 VAKNRVCL-PNRGFLKQLRELDKQLV 201
DB 287 VKORRGVISPFPNFSFMGLQLQLETVL 312
RESULT 7
DUS4_CHICK
ID DUS4_CHICK STANDARD; PRT; 375 AA.
AC Q9PW71;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)

20-AUG-2001 (Rel. 40, Last annotation update)
DUAL SPECIFICITY PROTEIN PHOSPHATASE 4 (EC 3.1.3.48) (EC 3.1.3.16)
(MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE-2) (MAP KINASE
PHOSPHATASE-2) (MKP-2).
DUSP4 OR MKP2.
Gallus gallus (Chicken).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
Gallus.
NCBI_TaxID=9031;
[1]
SEQUENCE FROM N.A.
STRAIN=white leghorn;
MEDLINE=20379359; PubMed=10918612;
Fu S.-L., Waba A., Vogt P.K.;
"Identification and characterization of genes upregulated in cells
transformed by v-Jun";
Oncogene 19:3537-3545(2000).
-!- FUNCTION: REGULATES MITOGENIC SIGNAL TRANSDUCTION BY
DEPHOSPHORYLATING BOTH THR AND TYR RESIDUES ON MAP KINASES
ERK1 AND ERK2 (BY SIMILARITY).
-!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATASE + H(2)O =
PROTEIN TYROSINE + ORTHOPHOSPHATE.
-!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
-!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AF167296; AAD46536.1; -
DR HSSP; Q16828; IMKP.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Nuclear protein.
FT DOMAIN 30 45 CH2 A DOMAIN.
FT DOMAIN 122 137 CH2 B DOMAIN.
FT DOMAIN 178 375 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 261 261 BY SIMILARITY
SQ SEQUENCE 375 AA; 41052 MW; 179290D0C2BEEF1 CRC64;

Query Match 15.5%; Score 180; DB 1; Length 375;
Best Local Similarity 30.1%; Pred. No. 3.9e-10;
Matches 55; Conservative 35; Mismatches 75; Indels 18; Gaps 6;

QY 20 LSPKMEEGEEDYCTPGAFELERLFWKSPQYTHVNEVMPKLYIGDEATALDRYLQKA 79
DB 151 ISPPSSAESLDLGFSCGT----PLNDQGGPV-----EILPFLYLGSAVHAARRDMLDAL 201
QY 80 GTHVLNAAHGWNVDTPDYRRMDIQYHGVEADDLPFDLSVFFYPAAAFIDRLSD 139
DB 202 GITALLNVS-----SDCPNHP-EGHYQYKCIPIVDNHNKADISSWFMEAEIYD-SVREC 253
QY 140 HSKILVHCVMGSRATLVLAYLMTKDMTLVDAIQQVAKNCVL-PNRGFLKQLRELDK 198
DB 254 CGRLVHCQAGISRSATTCLAYLMMKKRVKLEKAFEFVQKRSIIISPNFSGQLQLFES 313
QY 199 QLV 201
:::

Db 314 QVL 316
RESULT 8
DUS1_HUMAN
ID DUS1_HUMAN STANDARD; PRT; 367 AA.
AC P28562;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP KINASE PHOSPHATASE-1) (MKP-1) (PROTEIN-TYROSINE PHOSPHATASE
DE CL100) (DUAL SPECIFICITY PROTEIN PHOSPHATASE HVH1).
GN DUSP1 OR PTPN10 OR MKP1 OR CL100 OR VH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=foreskin;
RX MEDLINE=93024952; PubMed=1406996;
RA Keyes S.M., Emslie E.A.;
RT "Oxidative stress and heat shock induce a human gene encoding a
RT protein-tyrosine phosphatase.";
RL Nature 359:644-647(1992).
CC -!- FUNCTION: DUAL SPECIFICITY PHOSPHATASE THAT DEPHOSPHORYLATES MAP
CC KINASE ERK2 ON BOTH THR-183 AND TYR-185.
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -!- INDUCTION: BY OXIDATIVE STRESS AND HEAT SHOCK.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; X68277; CAA48338.1; -
DR PIR; S29090; S29090.
DR HSSP; P51452; 1VHR.
DR MIM; 600714; -
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Cell cycle.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 BY SIMILARITY
SQ SEQUENCE 367 AA; 39297 MW; 11BD1D39A9FCD51F CRC64;

Query Match 14.6%; Score 169; DB 1; Length 367;
Best Local Similarity 30.2%; Pred. No. 4.5e-09;
Matches 48; Conservative 33; Mismatches 64; Indels 14; Gaps 5;

QY 44 LEWKGSPQYTHVNEVMPKLYIGDEATALDRYLQKAGFTVHNAAGRWVDTGPDYRD 103
DB 168 LYDQGGPV-----EILPFLYLGSAVHASKOMLDALGITALINVS-----ANCPNHF-E 215
QY 104 MDIQYHGVEADDLPFDLSVFFYPAAAFIDRLSDHSHKILVHCVMGSRATLVLAYLM 163
DB 216 GHYQYKSIPIVDNHNKADISSWFNEAIDFID-SIKNAGGRVHVHCQAGISRSATTCIAYLM 274
:::

DR HSP; P51452; 1VHR.
DR MGD; MGI:105120; Ptpn16.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00556; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00554; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Cell cycle.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258
FT MUTAGEN 258 258 C->S; LOSS OF ACTIVITY.
SQ SEQUENCE 367 AA; 39369 MW; 50B5F90FEBB019AB CRC64;

Query Match 14.3%; Score 166; DB 1; Length 367;
Best Local Similarity 30.2%; Pred. No. 8.8e-09;
Matches 48; Conservative 32; Mismatches 65; Indels 14; Gaps 5;

QY 44 LFWKSPQYTHVNEVPKLYIGDEATDRLYRLQKAGFTHVLAAGRWNVDTGPDYRD 103
Db 168 LYDGGPV-----EILSFLYLGSAVHASKMDLALGITALINVS-----ANCPNHF-E 215
QY 104 MDIOYHGVEADDLPFDLSVFYPAAFIDRALSDHSHKILVHCVMGRSRSATLVLAYLM 163
Db 216 GHYQKSPVEDNHKADISSWFNEAIDFID-SIKDAGRVFVHCQGISRSATICLAYLM 274
QY 164 IHKMTLVDAIOQVAKRCVL-PNRGFLKQLRELDKQLV 201
Db 275 RTNRVKLDEAFEFVKQRSSIIISPNFEMGQLLQFESQVL 313

RESULT 13
DUS1_RAT
ID DUS1_RAT STANDARD; PRT; 367 AA.
AC Q64623;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP KINASE PHOSPHATASE-1) (MKP-1) (PROTEIN-TYROSINE PHOSPHATASE
CL100).
GN DUSP1 OR CL100.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Lung;
RA Muda M., Schlegel W., Arkinstall S.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: DUAL SPECIFICITY PHOSPHATASE THAT DEPHOSPHORYLATES MAP
KINASE ERK2 ON BOTH THR-183 AND TYR-185
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR HSP; P51452; 1VHR.

DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00556; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00554; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Cell cycle.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 BY SIMILARITY.
SQ SEQUENCE 367 AA; 39541 MW; 5112ADF290499139 CRC64;

Query Match 14.3%; Score 166; DB 1; Length 367;
Best Local Similarity 30.2%; Pred. No. 8.8e-09;
Matches 48; Conservative 32; Mismatches 65; Indels 14; Gaps 5;

QY 44 LFWKSPQYTHVNEVPKLYIGDEATDRLYRLQKAGFTHVLAAGRWNVDTGPDYRD 103
Db 168 LYDGGPV-----EILSFLYLGSAVHASKMDLALGITALINVS-----ANCPNHF-E 215
QY 104 MDIOYHGVEADDLPFDLSVFYPAAFIDRALSDHSHKILVHCVMGRSRSATLVLAYLM 163
Db 216 GHYQKSPVEDNHKADISSWFNEAIDFID-SIKDAGRVFVHCQGISRSATICLAYLM 274
QY 164 IHKMTLVDAIOQVAKRCVL-PNRGFLKQLRELDKQLV 201
Db 275 RTNRVKLDEAFEFVKQRSSIIISPNFEMGQLLQFESQVL 313

RESULT 14
DUS8_HUMAN
ID DUS8_HUMAN STANDARD; PRT; 625 AA.
AC Q13202;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE HVH-5).
GN DUSP8 OR VH5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96009533; PubMed=7561881;
RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.;
RT "HVH-5: a protein tyrosine phosphatase abundant in brain that
inactivates mitogen-activated protein kinase.";
RL J. Neurochem. 65:1823-1833(1995).
CC -!- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
SIMILARITY).
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -!- TISSUE SPECIFICITY: ABUNDANT IN BRAIN, HEART, AND SKELETAL MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U27193; AAA83151.1; -

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X95518; CAA64772.1; -
CC MGD; MGI; 106626; Ntclp1.
CC InterPro; IPR000340; DS_phosphatase.
CC InterPro; IPR001763; Rhodanese_domain.
CC InterPro; IPR000387; TYR_phosphatase.
CC Pfam; PF00782; DSPC; 1.
CC SMART; SM00195; DSPC; 1.
CC SMART; SM00450; RHOD; 1.
CC PROSITE; P500383; TYR_PHOSPHATASE_1; 1.
CC PROSITE; P50056; TYR_PHOSPHATASE_2; 1.
CC PROSITE; P50054; TYR_PHOSPHATASE_DUAL; 1.
CC KW Hydrolase; Nuclear protein.
CC DOMAIN 28 43 CH2 A DOMAIN.
CC FT DOMAIN 117 132 CH2 B DOMAIN.
CC FT DOMAIN 162 432 PROTEIN-TYROSINE PHOSPHATASE.
CC FT DOMAIN 452 459 POLY-ARG.
CC FT DOMAIN 555 558 POLY-SER.
CC FT DOMAIN 559 576 POLY-GLY.
CC FT DOMAIN 577 600 POLY-SER.
CC FT DOMAIN 311 552 PRO-RICH.
CC FT ACT_SITE 246 246 BY SIMILARITY.
CC SQ SEQUENCE 663 AA; 68847 MW; 416F429A12C1FA7C CRC64;

Query Match 14.0%; Score 163; DB 1; Length 663;
Best Local Similarity 31.8%; Pred. No. 3.7e-08;
Matches 49; Conservative 31; Mismatches 52; Indels 22; Gaps 7;

QY 55 VNEWPKLYIGDENTALDRYELQKAGFTHVLNAAHGRRNVDTGPDYRDMDIQYHGVEAD 114
: : ||| : : : : ||| : | || : : : : : : : : : : : : : : : : : :
Db 161 LTRILPHLYLGSQDVLNKLMDLMTQNGISYVLNAS---NSCPKPDFI--CESRFMRIPIN 214
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 115 D-----LPTFDLSVFYFAAAFTDRA--LSDHSHKILVHCVMGRSRSATLVLAYLMTHKD 167
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 DNYCEKLLPWLDKSI-----EFIDRAKLSS--CQVIVHCLAGISRSATIAIYIMKTMG 266
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 168 MTLVDAIQQVAKNR--CVLPNRRGFLKQLRELDKQL 200
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 267 MSSDDAYRFVKDRRPSISPNFNLGQLLEVERSL 300
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: February 8, 2002, 15:45:16
Job time: 190 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:40:56 ; Search time 15.19 seconds
(without alignments)
1103.252 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 1161

Sequence: 1 MTSGEVKTSLKNAYSAKRL.....VOQRRSQRQDGEEDGREL 220

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PTR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	346.5	29.8	185	A47196	dual specificity p
2	204.5	17.6	365	T32494	hypothetical prote
3	192.5	16.6	303	T46405	hypothetical prote
4	192.5	16.6	393	A56947	dual specificity p
5	192.5	16.6	394	A56115	dual specificity p
6	184.5	15.9	314	B57126	dual specificity p
7	180.5	15.5	314	A57126	dual specificity p
8	174	15.0	223	T49365	protein tyrosine p
9	169	14.6	367	T29090	dual specificity p
10	168	14.5	205	T49364	protein tyrosine p
11	168	14.5	226	T21380	dual specificity p
12	166	14.3	367	S24411	dual specificity p
13	166	14.3	367	S22655	dual specificity p
14	160.5	13.8	142	T03074	dual specificity p
15	159	13.7	283	G84458	probable protein p
16	158	13.6	330	T39698	protein tyrosine p
17	157.5	13.6	384	T13890	dual specificity p
18	150.5	13.0	186	T16056	hypothetical prote
19	137	11.8	272	T18915	hypothetical prote
20	136.5	11.8	276	T48906	protein-tyrosine-p
21	134.5	11.6	204	T17802	hypothetical prote
22	132.5	11.4	619	T15969	hypothetical prote
23	132	11.4	209	S48459	probable dual spec
24	124.5	10.7	489	S58725	dual specificity p
25	124	10.7	771	T47666	phosphatase-like p
26	122	10.5	364	S31304	protein-tyrosine-p
27	113	9.7	171	Q0VZHI	dual specificity p
28	111	9.6	292	S41012	hypothetical prote
29	110.5	9.5	169	T30684	probable dual spec

ALIGNMENTS

RESULT 1

A47196

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) DUSP3 [validated] - human
N:Alternate names: protein-tyrosine-phosphatase VHR; VHL-related dual-specificity ph
C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence_revision 01-Mar-1996 #text_change 15-Sep-2000

C:Accession: A47196; A58760

B:Shibasaki, T.; Bottaro, D.P.; Chan, A.; Mikl, T.; Aaronson, S.A.

Proc. Natl. Acad. Sci. U.S.A. 89, 12170-12174, 1992

A:Title: Expression cloning of a human dual-specificity phosphatase.

A:Reference number: A47196; MUID:93101689

A:Accession: A47196

A:Molecule type: mRNA

A:Residues: 1-185 <ISH>

A:Cross-references: GB:L05147; NID:gl81839; PIDN:AAA35777.1; PID:gl81840

A:Experimental source: fibroblasts

A:Note: sequence extracted from NCBI backbone (NCBIN:120790, NCBI:P:120791)

B:Denu, J.M.; Zhou, G.; Wu, L.; Zhao, R.; Yuanyan, J.; Saper, M.A.; Dixon, J.E.

J. Biol. Chem. 270, 3796-3803, 1995

A:Title: The purification and characterization of a human dual-specific protein tyro

A:Reference number: A58760; MUID:95181338

A:Accession: A58760

A:Molecule type: protein

A:Residues: 2-7 <DEN>

R:Yuanyan, J.; Denu, J.M.; Dixon, J.E.; Saper, M.A.

submitted to the Brookhaven Protein Data Bank, February 1996

A:Reference number: A66852; PDB:1VHR

A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 8-185

R:Zhou, G.; Denu, J.M.; Wu, L.; Dixon, J.E.

J. Biol. Chem. 269, 28084-28090, 1994

A:Title: The catalytic role of Cys(124) in the dual specificity phosphatase VHR.

A:Reference number: A55447; MUID:95050584

A:Contents: annotation; active site

C:Genetics:

A:Gene: GDB:DUSP3; VHR

A:Cross-references: GDB:342110; OMIM:600183

A:Map position: 17q21-17q21

C:Function:

A:Description: catalyzes hydrolysis of peptidyl-phosphoserine, -phosphothreonine and
C:Superfamily: dual specificity phosphoprotein phosphatase DUSP3; VHL-type dual spec

C:Keywords: phosphoprotein; phosphoric monoester hydrolase

F:2-185/Product: dual specificity phosphatase 3 #status predicted <MAT>

F:37-177/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>
F:124/Active site: Cys (phosphocysteine intermediate) #status experimental

F:130/Binding site: substrate phosphate (Arg) #status predicted

* Query Match

Best Local Similarity 29.8%; Score 346.5; DB 1; Length 185;

Matches 78; Conservative 35; Mismatches 54; Indels 11; Gaps 3;

Oy 37 GAFE-----LERLFVKSGPQYT-----HVNVEVMPKLYIGDEATLDRLQKAGFTHVLA 87

Db 3 GSFELSDQLNDLLSDGSCYSLPQPCNEVTPRIYVGNASVAQDIPKLGKLGITHVLNA 62
QY 88 ANGR--WNVDTGPDYRMDIOYHGVEADDLPFDLSVFFYPAAAFIDRALSDHSHKILV 145
Db 63 AGRSFMHVNTNANFYKDSGITYLGKANDTQEFNLSAYFERAADFIDQALAKNGRVLV 122
QY 146 HCVMGSRSATLVLAYLMIHKMTLVDAIQVAKNRCVLPNNGFLKQLRELDKQLVQO 203
Db 123 HCREGYSRPTLVIAIYLMWRKMDVKSALSIVRQNRREICPNDGFLAQLCOLNDRLAKE 180

RESULT 2

T32494
hypothetical protein C05B10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32494
R:Geisels, C.; Wamsley, P.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid C05B10.
A:Reference number: Z21178
A:Accession: T32494
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-365 <GET>
A:Cross-references: EMBL:AF036685; PIDN:AAB88308.1; GSPDB:GN00022; CESP:C05B10.1
A:Experimental source: strain Bristol N2; clone C05B10
C:Genetics:
A:Gene: CESP:C05B10.1
A:Map position: 4
A:Introps: 22/2; 46/2; 74/3; 163/2; 204/2; 248/1; 287/2

Query Match 17.6%; Score 204.5; DB 2; Length 365;
Best Local Similarity 38.6%; Pred. No. 3.4e-11;
Matches 54; Conservative 24; Mismatches 55; Indels 7; Gaps 3;

QY 62 LVIGDEATLDYRLQKAGFTHVLAHAGRWNVDTGPDYRMDIOYHGVEADDLPFDL 121
Db 188 LVLGNETAKNDVLUKKYSISIVNTSNLPNT-----FEEDPNMRYLRISADDNASHNL 242
QY 122 SVFFYPAAAFIDRALSDHSHKILVHCVMGSRSATLVLAYLMIHKMTLVDAIQVAK-N 180
Db 243 TRFFPEATISFIDARRND-SACLVHCLAGISRSVTICLAYLMKTECTLDSAYENVQRN 301
QY 181 RCVLPNNGFLKQLRELDKQL 200
Db 302 ASIAPNFHFMGQLTDYERKL 321

RESULT 3

T46405
hypothetical protein DKFZp434O1321.1 - human
C:Species: Homo sapiens (man)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
C:Accession: T46405
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23034
A:Accession: T46405
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-303 <AAA>
A:Cross-references: EMBL:AL137704
A:Experimental source: adult testis; clone DKFZp434O1321
C:Genetics:
A:Note: DKFZp434O1321.1
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHI-type dual specificity

Query Match 16.6%; Score 192.5; DB 2; Length 303;
Best Local Similarity 34.2%; Pred. No. 3.4e-10;

Matches 50; Conservative 29; Mismatches 58; Indels 9; Gaps 4;
QY 57 EWPVKLYIGDEATLDYRLQKAGFTHVLAHAGRWNVDTGPDYRMDIOYHGVEADDL 116
Db 107 EILPPLYLGSAYHAARRDMLDALGITALLNVS-----SDCPNHF-EGHYQYKCIPVEDN 159
QY 117 PTFDLSVFFYPAAAFIDRALSDHSHKILVHCVMGSRSATLVLAYLMIHKMTLVDAIQO 176
Db 160 HKADISSWFMEAEIYD-AVKDCRGVLVHCOAGISRSATICLAYLMKKRVRLEAEFEF 218
QY 177 VAKNRCVL-PNRGFLKQLRELDKQLV 201
Db 219 VKQRSIIISPNSFNGQLLFESQVL 244

RESULT 4

A56947
dual specificity phosphatase (EC 3.1.3.-) HVH2 - rat
N:Alternate names: mitogen-activated protein kinase phosphatase 2
C:Species: Rattus norvegicus (Norway rat)
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-1999
C:Accession: A56947
R:Misra-Press, A.; Rim, C.S.; Yao, H.; Roberson, M.S.; Stork, P.J.S.
J. Biol. Chem. 270, 14587-14596, 1995
A:Title: A novel mitogen-activated protein kinase phosphatase. Structure, expression
A:Reference number: A56947; MUID:95301550
A:Accession: A56947
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-393 <Mis>
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHI-type dual specificity
C:Keywords: phosphoprotein; phosphoric monoester hydrolase
E:202-333/Domain: VHI-type dual specificity phosphoprotein phosphatase homology <VHI
F:279/Active site: Cys (phosphocysteine intermediate) #status predicted
F:285/Binding site: substrate phosphate (Arg) #status predicted

Query Match 16.6%; Score 192.5; DB 2; Length 393;
Best Local Similarity 34.2%; Pred. No. 4.8e-10;
Matches 50; Conservative 29; Mismatches 58; Indels 9; Gaps 4;

QY 57 EWPVKLYIGDEATLDYRLQKAGFTHVLAHAGRWNVDTGPDYRMDIOYHGVEADDL 116
Db 197 EILPPLYLGSAYHAARRDMLDALGITALLNVS-----SDCPNHF-EGHYQYKCIPVEDN 249
QY 117 PTFDLSVFFYPAAAFIDRALSDHSHKILVHCVMGSRSATLVLAYLMIHKMTLVDAIQO 176
Db 250 HKADISSWFMEAEIYD-AVKDCRGVLVHCOAGISRSATICLAYLMKKRVRLEAEFEF 308
QY 177 VAKNRCVL-PNRGFLKQLRELDKQLV 201
Db 309 VKQRSIIISPNSFNGQLLFESQVL 334

RESULT 5

A56115
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 4 - human
N:Alternate names: dual specificity phosphatase HVH2
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Apr-1998
C:Accession: A56115
R:Guan, K.L.; Butch, F.
J. Biol. Chem. 270, 7197-7203, 1995
A:Title: Isolation and characterization of a novel dual specific phosphatase, HVH2,
A:Reference number: A56115; MUID:95221370
A:Accession: A56115
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-394 <GUA>
A:Cross-references: GB:U01108
C:Genetics:
A:Gene: GDB:DU5P4; HVH2; MKP-2
A:Cross-references: GDB:433893

A:Map position: 8p21-8p11.2
C:Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
C:Keywords: phosphoprotein; phosphoric monoester hydrolase
F:203-334/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
F:280/Active site: Cys (phosphocysteine intermediate) #status predicted
F:286/Binding site: substrate phosphate (Arg) #status predicted

Query Match 16.6%; Score 192.5; DB 2; Length 394;
Best Local Similarity 34.2%; Pred. No. 4.8e-10;
Matches 50; Conservative 29; Mismatches 58; Indels 9; Gaps 4;
QY 57 EVMPLKLYIGDEATADLRQLKAGTFHVLNAAGRWNVDTGPDYRDMDIQYHGVEADDL 116
DB 198 EILPLYLGSAYHARRDMLDALTALLNVS-----SDCPNHP-EGHYQYKCPVEDH 250
QY 117 PTFDLVSFFYPAAAFIDRALSDHSHKILVHCVMGSRSRATLVLAYLMHKMTLVDAIQ 176
DB 251 HKADISSWFMEAEYID-AVKDCRGRLVHVCQAGISRSATICLAYLMKKRVRLEAEFEF 309
QY 177 VAKNRCVL-PNRGFLKQLRELDKQLV 201
DB 310 VKQRSIIISPNFSGQLLOPESQVL 335

RESULT 6
B57126
Dual specificity phosphatase (EC 3.1.3.-) 2 - mouse
N:Alternate names: mitogen-induced nuclear protein-tyrosine-phosphatase PAC-1
C:Species: Mus musculus (house mouse)
C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 24-Apr-1998
C:Accession: B57126
R:Rohan, P.J.; Davis, P.; Moskaluk, C.A.; Kearns, M.; Krutzsch, H.; Siebenlist, U.; Kell
Science 259, 1763-1766, 1993
A:Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase
A:Reference number: B57126
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-314 <ROH>
A:Cross-references: GB:L11330
C:Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity
C:Keywords: nucleus; phosphoprotein; phosphoric monoester hydrolase
F:180-311/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
F:257/Active site: Cys (phosphocysteine intermediate) #status predicted
F:263/Binding site: substrate phosphate (Arg) #status predicted

Query Match 15.9%; Score 184.5; DB 2; Length 314;
Best Local Similarity 28.3%; Pred. No. 2e-09;
Matches 56; Conservative 39; Mismatches 84; Indels 19; Gaps 6;
QY 10 LKNAYSSAKRLSPKMEEGEEDYCTPCAFELER-----LFWKGSPOYTHVNEVWPKLYI 64
DB 128 LRGGKFSQYTCPCDLCSEAPAAQALPPAGAENSNDPRPIYDQGGPV-----EILPYLYL 182
QY 65 GDEATADLRQLKAGTFHVLNAAGRWNVDTGPDYRDMDIQYHGVEADDLPTFDLSVF 124
DB 183 GSCNHSDDLQGLQAGITAVLNVS-----ASCPNHFEGFL-PHYKSIPIVEDNQVWEISAW 235
QY 125 FYPAAFIDRALSDHSHKILVHCVMGSRSRATLVLAYLMHKMTLVDAIQOAVKNCVL 184
DB 236 FQEAISFTID-SVKNSGGRVLVHVCQAGISRSATICLAYLIQSHRVRLEAEFVKQRGVI 294
QY 185 -PNRGFLKQLRELDKQLV 201
DB 295 SPNFSFMGQLLOLETQVL 312

RESULT 7
A57126
Dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 2 - human
N:Alternate names: mitogen-induced nuclear tyrosine phosphatase; phosphatase of activate

C:Species: Homo sapiens (man)
C:Date: 03-Nov-1995 #sequence_revision 02-Jul-1996 #text_change 11-Jun-1999
C:Accession: A57126
R:Rohan, P.J.; Davis, P.; Moskaluk, C.A.; Kearns, M.; Krutzsch, H.; Siebenlist, U.;
Science 259, 1763-1766, 1993
A:Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
A:Reference number: A57126; MUID:93206122
A:Accession: A57126
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-314 <ROH>
A:Cross-references: GB:L11329; NID:9559539; PIDN:AAA50779.1; PID:g292376
C:Genetics:
A:Gene: GDB:DUSP2
A:Cross-references: GDB:L39200
A:Map position: 2q11-2q11
C:Function:
A:Description: catalyzes the hydrolysis of peptidyl-phosphoserine, -phosphothreonine
C:Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specific
C:Keywords: nucleus; phosphoprotein; phosphoric monoester hydrolase
F:140-160/Region: nuclear location signal
F:180-311/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
F:257/Active site: Cys (phosphocysteine intermediate) #status predicted
F:263/Binding site: substrate phosphate (Arg) #status predicted

Query Match 15.5%; Score 180.5; DB 1; Length 314;
Best Local Similarity 32.2%; Pred. No. 4.6e-09;
Matches 47; Conservative 32; Mismatches 58; Indels 9; Gaps 4;
QY 57 EVMPLKLYIGDEATADLRQLKAGTFHVLNAAGRWNVDTGPDYRDMDIQYHGVEADDL 116
DB 175 EILPLYLFGSCSHSSDLQGLQAGITAVLNVS-----ASCPNHFEGFL-FRYSKSIPIVEDN 227
QY 117 PTFDLVSFFYPAAAFIDRALSDHSHKILVHCVMGSRSRATLVLAYLMHKMTLVDAIQ 176
DB 228 QMVEISAWFQEAIGFID-WKNSGGRVLVHVCQAGISRSATICLAYLMQSRVRLEAEFDF 286
QY 177 VAKNRCVL-PNRGFLKQLRELDKQLV 201
DB 287 VKQRGVISPNFSGQLLOLETQVL 312

RESULT 8
I49365
Protein tyrosine phosphatase - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I49365
R:Wishart, M.J.; Denu, J.M.; Williams, J.A.; Dixon, J.E.
J. Biol. Chem. 270, 26782-26785, 1995
A:Title: A single mutation converts a novel-phosphotyrosine binding domain into a d
A:Reference number: I49364; MUID:96070766
A:Accession: I49365
A:Status: preliminary; translated from GB/EMBL/PDBJ
A:Molecule type: mRNA
A:Residues: 1-223 <RES>
A:Cross-references: EMBL:U34973; NID:gl063624; PIDN:AAA87037.1; PID:gl063626
C:Genetics: 168/3
A:Introns: 168/3
C:Superfamily: Vhl-type dual specificity phosphoprotein phosphatase homology
F:36-174/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>

Query Match 15.0%; Score 174; DB 2; Length 223;
Best Local Similarity 30.2%; Pred. No. 1.2e-08;
Matches 55; Conservative 31; Mismatches 68; Indels 28; Gaps 7;
QY 22 PKMEEGEEDYCTPCAFELERLFWKGSPOYTHVNEVWPKLYIGDEATADLR--LQKA 79
DB 12 PCKDDAEWY-----PMRMEQEVLPGLFLGYPYSAMSKLPILQKH 55
QY 80 GFTHVLNAAGRWNVDTG---PDYRDMDIQYHGVEADDLPTFDLSVFFYPAAAFIDRAL 136

Db 56 GITHIICI---RQNEANFIKPNF--QQLFRYLVDIADNPVENIIRFPMTKEFDGSL 110
QY 137 SDDHSHKILVHCVMGSRSSATLVLAIVLMIHKDMTLVDATQOQVAKNR-CVLPNRRGFLKQLRE 195
Db 111 QNG-GKVLVHGNAGISRSAAAFVIAYIMETFGMKYRDAFAYVOERRFCINPNAGFVHQLQE 169
QY 196 LD 197
Db 170 YE 171

RESULT 9
S29090

N:Alternate names: protein-tyrosine phosphatase (EC 3.1.3.-) 1 - human
C:Species: Homo sapiens (man)
C:Date: 25-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999
C:Accession: S29090; A53052
R:Keyse, S.M.; Emelie, E.A.
Nature 359, 644-647, 1992
A:Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine phosphatase
A:Reference number: S29090; MUID:93024952
A:Accession: S29090
A:Molecule type: mRNA
A:Residues: 1-367 <KEY>
A:Cross-references: EMBL:X68277; NID:g29980; PIDN:CAA48338.1; PID:g29981
R:Kwak, S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.
J. Biol. Chem. 269, 3596-3604, 1994
A:Title: Isolation and characterization of a human dual specificity protein-tyrosine phosphatase
A:Reference number: A53052; MUID:94148864
A:Accession: A53052
A:Molecule type: DNA
A:Residues: 1-367 <KWA>
A:Experimental source: leukocyte
A:Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804, NCBIN:143806)
C:Genetics:
A:Gene: GDB:DUSP1; PTPN10
A:Cross-references: GDB:136197; OMIM:600714
A:Map position: 5q34-5q34
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity phosphoprotein; heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced phosphatase
F:181-312/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F:258/Active site: Cys (phosphocysteine intermediate) #status predicted
F:264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.6%; Score 169; DB 1; Length 367;
Best Local Similarity 30.2%; Pred. No. 6.4e-08;
Matches 48; Conservative 33; Mismatches 64; Indels 14; Gaps 5;

QY 44 LFWKSPQYTHVNEVWPVKLYIGDEATLDRYLQKAGFTHVLNAAHGRWNVDTGPDYRD 103
Db 168 LYDGGPV-----ETLPFLYLGSAVHASKOMLDALGITALINVS-----ANCPNHF-E 215
QY 104 MDIOYHGVEADLPFTDLSVFFYPAAAFIDRALSDHSHKILVHCVMGSRSSATLVLAIVL 163
Db 216 GHYQKSPVDEHNHKAIDISSWNEAIDFID-SIKNAGRVFVHCQGISRSATICLAYLM 274
QY 164 IHKDMTLVDATQOQVAKNR-CVLPNRRGFLKQLRELDKQLV 201
Db 275 RTNRVKLDEAFEVQKRSIIISPNFSPMQLQLQFESQVL 313

RESULT 10
I49364

C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I49364
R:Wishart, M.J.; Denu, J.M.; Williams, J.A.; Dixon, J.E.
J. Biol. Chem. 270, 26782-26785, 1995
A:Title: A single mutation converts a novel-phosphotyrosine binding domain into a dual-s

A:Reference number: I49364; MUID:96070766
A:Accession: I49364
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-205 <RES>
A:Cross-references: EMBL:U34973; NID:g1063624; PIDN:AAA87036.1; PID:g1063625
C:Superfamily: VH1-type dual specificity phosphoprotein phosphatase homology
F:36-174/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

Query Match 14.5%; Score 168; DB 2; Length 205;
Best Local Similarity 30.2%; Pred. No. 3.8e-08;
Matches 54; Conservative 30; Mismatches 67; Indels 28; Gaps 7;

QY 22 PMKEEGEEDYCTPGAFELERLFWKSPQYTHVNEVWPVKLYIGDEATLDRYR--LQKA 79
Db 12 PQCKDADEWTY-----PMRMEQEVLPGLFGLPYSSAMSKSLILQKH 55
QY 80 GFTVHVNAAHGRWNVDTG---PDYVRMDIOYHGVEADLPFTDLSVFFYPAAAFIDRAL 136
Db 56 GITHIICI---RQNEANFIKPNF--QQLFRYLVDIADNPVENIIRFPMTKEFDGSL 110
QY 137 SDDHSHKILVHCVMGSRSSATLVLAIVLMIHKDMTLVDATQOQVAKNR-CVLPNRRGFLKQLR 194
Db 111 QNG-GKVLVHGNAGISRSAAAFVIAYIMETFGMKYRDAFAYVOERRFCINPNAGFVHQLQ 168

RESULT 11
T21380

hypothetical protein F26A3.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21380
R:McMurray, A.
submitted to the EMBL Data Library, August 1996
A:Reference number: T21380
A:Accession: T21380
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-226 <WIL>
A:Cross-references: EMBL:Z78419; PIDN:CAB01700.1; GSPDB:GNO00019; CESP:F26A3.4
A:Experimental source: clone F26A3
C:Genetics:
A:Gene: CESP:F26A3.4
A:Map position: 1
A:Introns: 117/2; 150/3; 186/3

Query Match 14.5%; Score 168; DB 2; Length 226;
Best Local Similarity 30.8%; Pred. No. 4.3e-08;
Matches 56; Conservative 35; Mismatches 73; Indels 18; Gaps 7;

QY 44 LFWKSPQYTHVNEVWPVKLYIGDEATLDRYLQKAGFTHVLNAAHGRWNVDTGPDYRD 103
Db 3 LSFVRNPEYAAAMSEIVPGLFCG-VSALSDEMKKKTHIINAT-----TEVPNLS 54
QY 104 M-DIOYHGVEADLPFTDLSVFFYPAAAFID---RALSDHSHKILVHCVMGSRSSATLVL 159
Db 55 LGDIQRTKLWLEDTP-----QTYIYPHLELQSDQIQALLADGGKVLVHCVAGVRSASICL 110
QY 160 AVLMIHKDMTLVDATQOQV-AKNCVLPNRRGFLKQLRELDKQLVQORRRSQDGEEDGR 218
Db 111 AFLKYRCRNLEAYHLMKSRWVRPNLGFWRQLIAYE-QNVRENAGSVRLVRDEAQP 169
QY 219 EL 220
Db 170 QL 171

RESULT 12
S24411

N:Alternate names: 3CH134 protein; protein-tyrosine-phosphatase erp, nonreceptor ty

